National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

T-DNA TAGGING *IN VIVO*: ISOLATION AND CHARACTERIZATION OF A PUTATIVE PLANT PROMOTER MEDIATING GENE EXPRESSION IN TOBACCO FLOWERS

BY

CHAO JIANG ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY

IN

PLANT MOLECULAR BIOLOGY AND BIOTECHNOLOGY

DEPARTMENT OF PLANT SCIENCE

EDMONTON, ALBERTA

FALL, 1994

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-95198-2

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: **CHAO JIANG**

TITLE OF THESIS: **T-DNA TAGGING *IN VIVO*: ISOLATION AND CHARACTERIZATION OF A PUTATIVE PLANT PROMOTER MEDIATING GENE EXPRESSION IN TOBACCO FLOWERS**

DEGREE: **DOCTOR OF PHILOSOPHY**

YEAR THIS DEGREE GRANTED: **FALL, 1994**

PERMISSION IS HEREBY GRANTED TO THE UNIVERSITY OF ALBERTA LIBRARY TO REPRODUCE SINGLE COPIES OF THIS THESIS AND TO LEND OR SELL SUCH COPIES FOR PRIVATE, SCHOLARLY OR SCIENTIFIC RESEARCH PURPOSES ONLY.

THE AUTHOR RESERVES ALL OTHER PUBLICATION AND OTHER RIGHTS IN ASSOCIATION WITH THE COPYRIGHT IN THE THESIS, AND EXCEPT AS HEREINBEFORE PROVIDED NEITHER THE THESIS NOR ANY SUBSTANTIAL PORTION THEREOF MAY BE PRINTED OR OTHERWISE REPRODUCED IN ANY MATERIAL FROM WATHEVER WITHOUT THE AUTHOR'S PRIOR WRITTEN PERMISSION.

(SIGNED)

PERMANENT ADDRESS:

Harbin, Helongjiang

The People's Republic of China

DATE: May 18/94

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

THE UNDERSIGNED CERTIFY THAT THEY HAVE READ, AND RECOMMEND TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH FOR ACCEPTANCE, A THESIS ENTITLED **T-DNA TAGGING *IN VIVO*: ISOLATION AND CHARACTERIZATION OF A PUTATIVE PLANT PROMOTER MEDIATING GENE EXPRESSION IN TOBACCO FLOWERS** SUBMITTED BY **CHAO JIANG** IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF **DOCTOR OF PHILOSOPHY** IN **PLANT MOLECULAR BIOLOGY.**

Dr. A. Flanagan (Supervisor)

Dr. C. Hiruki

Dr. F. Yeh

Dr. D. Pilgrim

Dr. L. Comai (External)

Date: May 10/94

Dedication

This thesis is dedicated to the members of my family who made it possible for me to go through most of the emotional and financial hardships inherent to a graduate student's life: my son, Robert and my wife, Su Fan, for all their love, support and understanding during my graduate years, and to my parents, for their encouragement for me to pursue this level of education and especially to my grandmother, for her love in early years of my life.

ABSTRACT

A sensitive and nondestructive method is required to study the regulation of gene expression and T-DNA insertional mutagenesis in higher plants. In an attempt to accomplish this goal, a bacterial luciferase gene-based binary plant transformation vector was modified to construct two new vectors with enhanced capacity for cloning and expression of a foreign gene in higher plants. Furthermore, a promoterless bacterial luciferase *luxA* gene from *Vibrio harveyi* was used to construct a specialized T-DNA insertional vector to monitor the activation of plant genes *in vivo* and to study the regulation of gene expression in a native environment. Using this strategy, large numbers of transformed tobacco plants (253) were created by *Agrobacterium*-mediated transformation containing the plant insertional vector pPCVG *luxA&B*. Following *in vivo* screening, insertional mutants (45) were obtained that expressed the luciferase marker gene during different stages of plant growth. Of these, one transgenic plant that expressed the luciferase marker gene predominantly in flowers was identified and studied in detail. A putative promoter element which controls the organ specific gene expression of the luciferase was identified and rescued directly using inverse PCR. DNA sequence analysis of this fragment revealed the location of a TATA-like box and three CAAT boxes in the region closely linked to the 5' end of the promoterless *lux A* gene. In addition, a consensus sequence similar to the control regions found in other flower specific genes was identified. When this putative promoter region of 1065 bp was fused with the *lux F* marker gene and reintroduced by transformation into tobacco, the expression of the promoter-luciferase marker gene fusion was detected in the flowers. In conclusion, the luciferase gene-based binary plant transformation vectors offer a convenient tool for cloning and expression of a foreign gene in higher plants. Furthermore, an *in vivo* T-DNA tagging system provides a simple, sensitive and realistic approach for identification and isolation of transcriptional control regions of plant genes.

Acknowledgements

It is my great pleasure to thank my supervisor, Dr. A. Flanagan, for her support, guidance and encouragement during the writing of this thesis.

I would like specially to acknowledge my supervisory committee members, Dr. C. Hiruki and Dr. F. Yeh, for their support, helpful advice throughout the project and constructive criticism of this thesis.

I also would like to thank Dr. L. Comai and Dr. J. Singh for their critical reading of the thesis and to thank Dr. K.G. Briggs and Dr. J.A. Kernahan for their encouragement and financial support at the final stage of the writing of this thesis.

I wish to acknowledge my former supervisors, Dr. A.A. Szalay and Dr. W.H.R. Langridge, for the financial support through their research grant during this investigation.

I want to express my gratitude to Dr. M. Baga, Dr. M. Aguilar, Dr. A. Escher and Dr. R. Mayerhofer for their helpful discussions and advice at various stage of this investigation.

Finally, I would like to thank my wife, Su Fan, for her support, encouragement and patience throughout my graduate program.

TABLE OF CONTENTS

CHAPTER.....PAGE

I. GENERAL INTRODUCTION.....1

II. LITERATURE REVIEW.....6

1. *Agrobacterium* research.....7

1. A. Historical perspective.....7

1. B. Mechanism of T-DNA transfer.....9

2. Development of Ti plasmid-based Vectors.....11

3. Development of marker genes in higher plants.....13

4. Regulation of plant gene expression.....15

4. A. Transcriptional regulation of gene expression.....16

4.A.1. *Cis*-acting elements.....16

4.A.2. Methods for dissection of *cis*-acting elements.....20

4.B. Post-transcriptional regulation of gene expression.....24

5. T-DNA-mediated gene tagging.....27

5.A. Historical background of T-DNA-mediated gene tagging....27

5.B.Mechanics and strategies of T-DNA-mediated gene tagging.34

5. B. 1. Structure and organization of the integrated T-DNA.....34

5.B.2. Models of T-DNA intergration.....36

5.B.3. Strategies: T-DNA insertional vectors.....37

5.C. Regulation of transgene expression.....38

5.C.1. Copy number of T-DNA insert and transgene expression.....38

5.C.2. Position effect.....39

5.D. Phenotype and genetic analysis of T-DNA inheritance.....41

CHAPTER .....PAGE

5.D.1. Predicted phenotypes and genotypes of primary transformants.....41

5.D.2. T-DNA insertional mutagenesis in diploid and amphidiploid species.....42

6. Applications of T-DNA tagging in higher plants.....45

7. Genetic control of flower development.....48

7. Bibliography.....50

III. ALTERATION OF A LUCIFERASE BASED-BINARY PLANT TRANSFORMATION VECTOR FOR SIMULTANEOUS EXPRESSION OF DESIRED GENE PRODUCTS AND BACTERIAL LUCIFERASE IN TRANSGENIC PLANTS 76

1. Introduction.....77

2. Materials and methods.....78

Plasmids and strains.....78

Mobilization of binary plant expression vectors from *E. coli* into *A. tumefaciens*.....78

Isolation of binary plant expression vectors from *Agrobacterium* 79

Plant marterials and transformation.....80

Luciferase assays.....80

3. Results and discussion.....81

Modification of the *fab2* gene.....81

Addition of a synthetic polylinker.....82

Luciferase activity assay.....82

CHAPTER ..... PAGE

4. Bibliography 84

IV. IDENTIFICATION OF ACTIVATION PLANT GENES *IN VIVO* BY TAGGING WITH T-DNA BORDER-LINKED LUCIFERASE GENES 93

1. Introduction 94
2. Materials and methods 95
   - Bacterial strains and DNA manipulation 95
   - *Agrobacterium*-mediated transformation of tobacco leaf explants and screening of regenerated transgenic plants 95
   - *In vivo* detection of insertional mutant plants 96
   - Isolation of genomic DNA from transgenic plants 97
   - Identification of tobacco DNA fragments linked to the luciferase gene by Southern hybridization 98
   - Isolation and self-ligation of *lux*-containing genomic DNA fragments from transgenic plants for inverse PCR 99
   - Conditions for inverse PCR 100
3. Results and discussion 101
   - Construction of a specialized plant insertional vector and experimental strategy 101
   - Selection of insertional mutants based on the luciferase gene expression *in vivo* during plant development 102
   - Localization of the T-DNA insert by Southern hybridization 104
   - Amplification of plant genomic DNA flanking the T-DNA insertion by inverse PCR 104

CHAPTER ..... PAGE

Additional questions related to this chapter

Is T-DNA integration random or non-random?..... 106

The rationale for choosing amphodiploid tobacco..... 107

4. Bibliography..... 109

V. MOLECULAR CHARACTERIZATION OF A PUTATIVE PLANT PROMOTER RESPONSIBLE FOR MEDIATING GENE EXPRESSION IN TOBACCO FLOWERS..... 130

1. Introduction..... 131

2. Materials and methods..... 132

Bacterial strains and DNA manipulation..... 132

Plant transformation and regeneration 133

Screening transformed plants..... 133

Subcloning and DNA sequencing..... 134

3. Results and discussion..... 134

Subclones and sequencing strategies..... 134

Sequence analysis of a putative plant promoter..... 135

Consensus sequences..... 136

Construction of a promoter test vector, pTV-1p..... 138

The putative promoter mediated luciferase gene expression in transgenic Plants..... 139

Additional questions related to this chapter

Possible reasons for the decreased expression level observed with the cloned promoter..... 140

What additional evidence would be needed to prove that

CHAPTER ..... PAGE

the cloned DNA is indeed a promoter? 142

4. Bibliography 145

VI. GENERAL DISCUSSION AND FUTURE DIRECTION 169

1. General discussion 170

2. Future perspectives 172

3. Bibliography 175

LIST OF TABLE AND CHARTS

TABLE/CHART.....PAGE

TABLE IV-1. Distribution of luciferase gene expression in transgenic tobacco plants containing the plant insertional vector pPCVG *luxA&B*. 121

CHART IV-1. Differential expression of promoterless bacterial luciferase marker gene in transgenic tobacco plants.....123

CHART V-1. The FEP directed the *lux* gene expression in floral organs of transgenic plants. 168

LIST OF FIGURES

FIGURE.....PAGE

Fig. II-1. Diagram of ends of a transcript.....18

Fig. III-1. Diagram depicting the strategy employed in the construction of vectors pPCV701 *lux FM1* and pPCV701 *lux FM2*.....88

Fig. III-2. DNA sequence of the synthetic multiple cloning site (MCS).....90

Fig. III-3. Collected photon image of light emission *in vivo* from *A. tumefaciens* strain GV3101 pMP90RK containing vectors pPCV 701 *lux FM1* and pPCV701 *lux FM2*.....92

Fig. IV-1. Circular map of the plant insertional vector pPCVG *luxA&B*.....115

Fig. IV-2. Schematic representation of insertional activation of a promoterless bacterial luciferase marker gene in the plant genome.....117

Fig. IV-3. Bacterial luciferase gene expression visualized in different organs of transgenic tobacco plants detected by low light image analysis.....119

Fig. IV-4. Southern blot analysis of the plant 12-1 containing plant insertional vector pPCVG *luxA&B* with ^{32}P-radio-labeled *lux A* DNA fragment as a probe.....125

Fig. IV-5. Schematic representation of promoter isolation from transgenic plants based on T-DNA tagging with the promoterless luciferase gene fusion and inverse PCR methods.....127

Fig. IV-6. Direct amplification of the flanking plant DNA fragment

linked with the T-DNA tag from genomic DNA isolated

FIGURE.....PAGE

from the plant 12-1 by inverse PCR.....129

Fig. V-1. Bacterial luciferase expression visualized in flower of insertional mutant plant 12-1 detected by low-light vdeo image analysis.....154

Fig. V-2. Restriction enzyme analysis of two independent clones pBS-1 and pBS-2 containing an insert from inverse PCR.....156

Fig. V-3. Generation of undirectional deletions with exonuclease III.....158

Fig. V-4. Schematic representation of sequencing strategies of two independent clones pBS-1 (a) and pBS-2 (b) containing 1.384 kb insert from inverse PCR.....160

Fig. V-5. Structure of the PCR fragment and DNA sequence analysis of the FEP element.....162

Fig. V-6. Comparison of the T-DNA right border sequence before and after the integration.....164

Fig. V-7. Construction of a plant promoter test vector, pTV-1p, containing the FEP-*lux* F fusion.....166

LIST OF ABBREVIATIONS

aph(3')-II	Aminoglycoside phosphotransferase II gene
BAP	Benzylamino purine
bp	Base pair
BSA	Bovine serum albumin
CaMV	Cauliflower mosaic virus
CAT	Chloramphenicol acetytransferase
cm	Centimeter
CTAB	Cetyltriethylammonium bromide
CAAT box	Part of a conserved sequence located upstream of a gene
dNTPs	Deoxyribonucleotides (dATP, dCTP, dGTP, dTTP)
EDTA	Ethylene diaminetetraacetic acid, disodium salt
Exo III	Exonuclease III
FEP	Flower expression promoter
GUS	β-glucuronidase
hpt	Hygromycin phosphotransferase
hr(s)	Hour(s)
iPCR	Inverse polymerase chain reaction
kb	Kilobase
l	Liter
LB	Left border of the T-DNA
LUX	Luciferase
LU	Light unit
luxA	The gene encoding α subunit of luciferase
luxB	The gene encoding β subunit of luciferase

luxF	Luciferase αβ fusion gene
M	Molar
MCS	Multiple cloning sites
mg	Milligram
mg^{-1}	Per milligram
min	Minute
ml	Milliliter
mm	Millimeter
mM	Millimolar
MS medium	Murashige and Skoog medium
NAA	Naphthalene acetic acid
npt II	Neomycin phosphotransferase
nt	Nucleotide(s)
pmol	Picomolar
pol II	RNA polymerase II
RB	Right border of the T-DNA
RFLP	Restriction Fragment Length Polymorphism
SDS	Sodium dodecyl sulfate
sec^{-1}	Per second
20 x SSC	3.0 M NaCl, 0.3 M Na-citrate, pH 7.0
TAAT box	A conserved AT-rich sequence located in promoter
1 x TBE	89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA
T-DNA	Transfer DNA
1 x TE	10 mM Tris-HCl, 1 mM EDTA
Ti	Tumor inducing
U	Unit
µg	Microgram

µl	Microliter

Chapter I

General Introduction

Agrobacterium tumefaciens-mediated gene transfer is a commonly used method to introduce genes into plant cells, to study gene regulation and function in regenerated transformants and to isolate developmentally regulated genes or promoters from higher plants (Schell, 1987; Benfey and Chua, 1989; Gasser and Fraley, 1989; Walden *et al.*, 1991; Walbot, 1992). Therefore, development and improvement of novel plant transformation and T-DNA tagging vector systems not only offer a convenient tool for studying gene structure, function and regulation in higher plants, but also provide opportunities for identification and isolation of developmentally regulated promoter elements from higher plants. In the present study, the bacterial luciferase marker gene system was used to modify a binary plant transformation vector and to develop an *in vivo* T-DNA tagging system in higher plants.

Conventional approaches for isolation of promoters and genes involve differential screening of cDNA libraries made from mRNA isolated from different organs, and synthesis of oligonucleotides from protein sequence data, following protein purification or *in vitro* translation of isolated mRNAs. These approaches can only isolate genes encoding relatively abundant transcripts or proteins. Therefore, an alternative approach is needed for many important genes which are expressed at low levels.

T-DNA tagging is a recently developed method for promoter and gene isolation (Andre *et al.*, 1986; Koncz *et al.*, 1989). In this, T-DNA is transferred into plants by *Agrobacterium*-mediated gene transfer and is integrated into the plant genome by illegitimate recombination events (Mayerhofer *et al.*, 1991; Gheysen *et al.*, 1991). T-DNA (transferred DNA) resides in the Ti plasmid of *Agrobacterium tumefaciens* and is known to be a unique insertion element that is capable of being integrated into the plant nuclear genome

(Zambryski *et al.*. 1989) in or near the genes (Koncz *et al.*, 1989). Once inserted, T-DNA is a molecular tag for the insertion site. Therefore, T-DNA serves both as a mutagen via interruption of a plant gene and as a molecular tag for the isolation of flanking plant DNA. Some developmentally interesting genes, which may not necessarily highly efficient in encoding abundant transcripts or proteins, have been identified by T-DNA tagging (Andre *et al.*, 1986; Koncz *et al.*, 1989; Feldmann, 1991). Since it does not require abundant gene products or genetic information, T-DNA tagging is widely used to isolate developmentally regulated promoters and genes from higher plants (Walden *et al.*, 1991; Feldmann, 1991).

Once promoters are isolated, studies of their regulation have traditionally utilized promoter-reporter gene fusions *in vitro* to test function *in vivo* following *Agrobacterium*-mediated gene transfer. One obvious problem with this approach is that it is difficult to obtain consistent results from the same construct because of position and copy number effects (Dean *et al.*, 1988; Hobbs *et al.*, 1990; 1993). Furthermore, the size of an engineered construct is limited by the capability of current delivery systems. As such, eukaryotic genes containing regulatory elements in distant upstream or downstream cannot be studied by promoter-reporter gene fusion (Forget, 1993).

A more recent strategy, which has been used to study gene expression and isolation in *Drosophila* (Bellen *et al.*, 1989), *C. elegans* (Hope, 1991), mouse (Allen *et al.*,1988) and plants (Koncz *et al.*, 1989; Walden et al., 1991), relies on the activation of reporter genes such as neomycin phosphotransferase (*nptII*) (Teeri *et al.*, 1986), aminoglycoside phosphotransferase (*aph(3')-II*) (Andre *et al.*, 1986; Koncz *et al.*, 1989) and β-glucuronidase (GUS) (Kertbundit *et al.*, 1991; Topping *et al.*, 1991) by native

gene regulatory sequences. These reporter genes are either promoterless or contain a minimal promoter or without a ATG start codon. The rational behind the technique is that a reporter gene linked to the right or left border of the T-DNA may be transcriptionally activated by integration downstream of a native gene promoter to create either transcriptional or translational fusion. It is expected that the pattern of a reporter gene expression would reflect the pattern of expression of the native gene, and indeed, this has been demonstrated in both *Arabidopsis* and tobacco (Koncz *et al.*, 1989; Topping *et al.*, 1991; Kertbundit *et al.*, 1991). However, using *aph(3')-II* or GUS as a reporter gene in this approach has one drawback. Determination of gene expression in transgenic plants requires destructive analysis of plant tissues (Koncz *et al.*, 1992).

Ideally, monitoring the activation of a plant gene *in vivo*, studying the regulation of its expression in native environment during organogenesis and isolating promoters and their corresponding genes should be nondestructive. In an attempt to accomplish this goal, a promoterless bacterial luciferase *luxA* gene from a marine bacterium *Vibrio harveyi* (Belas *et al.*, 1982; Baldwin *et al.*, 1984; Cohn *et al.*, 1985) was used to construct a specialized T-DNA insertional vector. The bacterial luciferase marker gene offers a reproducible and sensitive *in vivo* assay system which has been widely used to study the regulation of gene expression both in prokaryotic and eukaryotic organisms (Koncz *et al.*, 1990; Langridge *et al.*, 1991). The bacterial luciferase marker gene in the present study was linked to the right border of a plant insertional vector pPCVG *luxA&B* which is a part of a functional insertional element. As such, it is integrated into the plant genome, and results in *in vivo* activation of the luciferase marker gene by a native gene regulatory sequence (Jiang *et al.*, 1992; 1993b). The strength of this

approach is that the activation of the luciferase gene may reflect the pattern of native gene expression. Therefore, the identification of insertional mutants can be based on the marker gene expression rather than on the mutant phenotype. Expression of the *lux* marker gene can be studied *in vivo* by low-light video image analysis (Jiang *et al.*, 1992). The bacterial luciferase marker gene can be used as an indicator for detection of a gene activation *in vivo*, and as a molecular tag for isolation of promoters and their corresponding genes (Jiang *et al.*, 1992).

The objectives of this research are as follows:

1. Alteration of a luciferase-based binary plant transformation vector for simultaneous expression of desirable gene products and bacterial luciferase in transgenic plants (Chapter III).
2. Utilization of a bacterial luciferase-based plant insertional vector system to create a variety of insertional mutant plants as sources for isolation of developmentally regulated plant promoters (Chapter IV).
3. Establishment of a detection system for monitoring the activation of a plant gene *in vivo* and selection of insertional mutants based on the expression of the luciferase gene (Chapter IV).
4. Establishment of a strategy to facilitate quick and direct cloning of putative promoters from plant genomes (Chapter IV).
5. Isolation and molecular characterization of a developmentally regulated plant promoter from plant genome (Chapter V).

Chapter II

Literature Review

In this chapter, the literature on historical background of *Agrobacterium* research, development of Ti-plasmid-based vectors and the use of marker genes in higher plants is first reviewed. This followed by a review of our current knowledge of mechanism of T-DNA-mediated gene tagging and related background of plant gene structure, transcriptional and translational control of gene expression. Strategies employed in creation of insertional mutant plants are also compared.

1. *AGROBACTERIUM* RESEARCH

1.A. Historical perspective

Agrobacterium is a plant pathogenic soil bacterium that belongs to the family *Rhizobiaceae*. The best known members of the *Agrobacterium* are *A. tumefaciens* and *A. rhizogenes*, which are, respectively, the causal agents of crown-gall and hairy-roots diseases. *Agrobacterium* has an extremely wide host range including most dicotyledonous and a few monocotyledonous plants (De Cleene and De Ley, 1976). Smith and Townsend first presented evidence that the bacterium which is now called *Agrobacterium tumefaciens* was the causative agent of the widespread neoplastic plant disease, crown gall (Smith and Townsend, 1907). Since then a large number of scientists throughout the world have focused their research on this organism in an effort to analyze the molecular mechanism underlying the process of crown gall induction in detail. This was driven by the hope that this would lead to a better understanding of oncogenesis in general, and to the development of remedies for such diseases. After a period of diminished interest in the system, *Agrobacterium* and crown gall research was revived when it became apparent that oncogenic gene transfer from *Agrobacterium* to plants might

form the molecular basis of crown gall induction, and thus the transfer system might be exploited for the genetic engineering of plants.

A key discovery, made almost 20 years ago, was the finding that virulent strains of *A. tumefaciens* contain a large extrachromosomal element harboring genes involved in crown gall induction (Zaenen *et al.*, 1974). This large extrachromosomal element turned out to be a plasmid of exceptionally large size (more than 200 kb). Because of its role in plant tumor induction, this plasmid was called the Ti (tumor-inducing) plasmid (van Larebeck *et al.*, 1975). The introduction of the Ti plasmid into related bacterial species such as the root nodule-inducing bacterium *Rhizobium trifolii* (Hooykaas *et al.*, 1977) or the leaf nodule inducer *Phyllobacterium myrsinacearium* (van Veen *et al.*, 1988) led to tumor-inducing strains, stressing the importance of the virulence determinants on the plasmid for the tumourigenicity of their bacterial hosts. However, introduction and maintenance in more distantly related bacteria such as *Escherichia coli* or *Pseudomonas aeruginosa* did not result in tumor-inducing strains (Hille *et al.*, 1983), indicating that other factors were also important.

In 1983, researchers directed by Van Montagu and Schell at the University of Gent (Belgium) and by Fraley at the Monsanto Co. (St. Louis, MO), independently reported that the crown gall disease-causing genes from the T-DNA had been removed (Zambryski *et al.*, 1983; Fraley *et al.*, 1983; Herrera-Estrella *et al.*, 1983b). Since then, dramatic progress has been made in the development of gene transfer systems for higher plants (Horsch *et al.*, 1985; De Block *et al.*, 1984), and *A. tumefaciens* has proven to be an efficient, highly versatile vehicle for the introduction of foreign genes into plants and plant cells (Gasser and Fraley, 1989).

Currently, *Agrobacterium*-mediated transformation is the most frequently used method to introduce foreign DNA into plant cells and to study gene structure, function and expression. The success rate of transformation is variable, and depends on the different species, cultivars of the host and methods employed. Nevertheless, *Agrobacterium*-mediated transformation has produced a great number of transgenic plants (Gasser and Fraley, 1989).

1.B. Mechanism of T-DNA transfer

Three genetic elements in *Agrobacterium* are essential for the transfer of T-DNA from Ti plasmids to plant cells: *vir* genes (Stachel and Nester, 1986), T-DNA border sequences (Yadav *et al.*, 1982; Zambryski, 1992) located on Ti plasmid, and the chromosomal virulence genes (Douglas *et al.*, 1985; Schell, 1987). The T-DNA is the mobile element, but it is not a classical transposable element, since it does not encode the products that mediate its movement. Instead, the Ti plasmid virulence region provides most of the trans-acting products for T-DNA transfer. The T-DNA is defined and delimited by two 25 bp direct repeats at its ends, the T-DNA borders. Any DNA between these borders is transferred by *Agrobacterium* into plant cells. Plasmids responsible for tumor-induction by *A. tumefaciens* and hairy-root induction by *A. rhizogenes* are termed Ti and Ri plasmids, respectively. These plasmids are transferred from *Agrobacterium* to plant cells during the infection process. The T-DNA is transported to the nucleus of infected plant cells and stably integrated into their nuclear genome (Schell, 1987).

T-DNA of Ti and Ri plasmids encode genes which are expressed in plants. Some of these genes are responsible for the synthesis or modification of basic plant growth substances, such as auxins and cytokinins. Regulation of these T-DNA genes alters the normal differentiation of plant cells and leads

to neoplastic growth or root proliferation. Other T-DNA genes direct the synthesis of particular amino acid and sugar derivatives, termed opines, which are specifically recognized and catabolized by *Agrobacteria* carrying Ti and Ri plasmids. Genes encoded by the T-DNA are not required for T-DNA transfer and integration. Rather, the process of T-DNA transfer is mediated by bacterial genes located in the virulence region of Ti and Ri plasmids. Mutations in this region generally lower the virulence, i.e. DNA transforming potential.

The *vir* region is 30 kb and organized into seven complementation groups, *virA*,,*virB*, *virC*, *virD*, *virE*, *virG*, and *virH* (Zambryski, 1992). The *virA* and *virG* loci encode positive regulatory proteins which directs *vir* gene expression (Stachel and Zambryski, 1986). The VirA product specifies an inner membrane protein which recognizes and responds to the presence of plant phenolic compounds. VirA transfers this information by protein phosphorylation to the product of *vir*G. VirG then acts as a transcriptional activator of itself and the other *vir* loci. The products of the *virC* and *virD* loci are enzymes which are involved in the generation and processing of the T-DNA copy. The products of *virB* and *virE* loci are involved in forming most of the structural components for facilitating T-DNA movement. The *virH* locus is not essential for virulence, but it may be involved in pathogenicity in specific dicotyledonous plant hosts (Kanemoto *et al.*, 1989). These virulence genes are regulated by a complex chemical sensory-regulatory system which is able to recognize plant metabolites produced during the synthesis of plant cell wall lignins. Wounds, therefore, not only provide physical entry for *Agrobacterium*, but also specifically induce the expression of bacterial virulence genes via lignin biosynthesis (Boulton *et al.*, 1986; Stachel *et al.*, 1985; Zambryski, 1992). Following induction of *vir* gene

expression, molecular reactions occur on the T-DNA element of the Ti plasmid to generate a transferable T-DNA copy. The virulence gene products mobilize the T-DNA by recognition and processing the 25 bp imperfect direct repeats located at the ends of Ti and Ri plasmid T-DNAs. The T-DNA processing leads to single- and double-stranded intermediates (Howard *et al.*, 1992). Of these, the single-stranded product, termed the T-strand, is expected to be transferred to the nucleus by a *vir* protein, which remains covalently bound to its 5'-end (Howard *et al.*, 1989). This protein carries a nuclear targeting signal and probably plays a role in the integration of the T-DNA by illegitimate recombination into the plant genome (Mayerhofer *et al.*, 1991; Gheysen *et al.*, 1991).

The interaction of *Agrobacterium* with plant cells is the only known natural example of inter-kingdom DNA transfer. Application of *Agrobacterium*-mediated transformation in plant molecular biology is based on the same natural process of tumor formation in wounded dicotyledonous plants by the soil bacterium, *A. tumefaciens* (Kado, 1991; Kahl and Schell, 1982; Ream, 1990; Zambryski *et al.*, 1989).

2. DEVELOPMENT OF TI PLASMID-BASED VECTORS

Several features of the T-DNA transfer process have been exploited in the development of vectors for both gene transfer and tagging (Gasser and Fraley, 1989; Schell, 1987; Walden *et al.*, 1991). First, the *vir* region functions in *trans*, thereby allowing the construction of a wide variety of transformation vectors which can be structurally independent from the Ti plasmid. Secondly, only the borders of T-DNA are required in *cis* for DNA transfer. Finally, the genes encoded by the T-DNA can be replaced without interfering with the transfer process. Therefore, it is possible to create a

disarmed vector by deleting the internal oncogenic region and replacing it with other DNA sequences. In early work, a 'disarmed' Ti plasmid vector, pGV3850, was constructed in which all of internal oncogenic regions (T-region) were replaced by the sequences from plasmid pBR322. The pBR322 DNA provides homology for cointegration of the T-region of pGV3850 with any pBR322 vector derivative carrying a gene of interest (De Block *et al.*, 1984). A triparental bacterial conjugation system was developed to transfer any pBR322 derivatives from *E. coli* to *A. tumefacins* pGV3850 by a single recombination step and to select for coconjugated pGV3850 containing the desired cloned genes (Van Haute *et al.*, 1983). Plant cells transformed with these modified plasmids were regenerated into whole plants without the disease symptom (Caplan *et al.*, 1983; Zambryski *et al.*, 1983). More recently, binary vectors having *vir* and T-DNA regions separated on different plasmids were developed (Hoekema *et al.*, 1983). The results from these experiments showed that the Ti-plasmid-encoded *vir* function could be located on a different plasmid and even on the bacterial chromosomal for T-DNA transfer (Hoekema *et al.*, 1983; 1984). Other improvements of Ti plasmid vectors were carried out by adding bacterial antibiotic resistance genes to allow selection of transformants on media containing the appropriate antibiotic (Bevan *et al.*, 1983; Barton *et al.*, 1983; Fraley *et al.*, 1983; Herrera-Estrella *et al.*, 1983a; 1983b). These include the Tn5-derived *npt II* gene which provides a high level of kanamycin resistance and the hygromycin phosphotransferase (*hpt*) gene which provides a high level of hygromycin resistance (Reiss *et al.*, 1984; Waldron *et al.*, 1985).

Disarmed Ti plasmids have been used successfully as vehicles for introduction of foreign genes into plant cells since the first transgenic plant was produced about ten years ago. The introduction of a foreign gene into

certain species of plants via the *Agrobacterium* system is now a common practice. The *Agrobacterium* system is being used extensively for the transfer of various traits to (crop) plants as well as for the study of function and regulation of gene expression in plants (Hooykaas and Schilperoot, 1992). Applications include the introduction of genes affecting widely diverse traits such as resistance to viruses (Abel *et al.*, 1986), herbicide tolerance (De Block *et al.*, 1987), altered flower color (van der Krol *et al.*, 1988), altered maturity of tomato (Smith *et al.*, 1988), male sterility (Mariani *et al.*, 1990), cold tolerance (Hightower *et al.*, 1991), altered source-sink relationships (Von Schaewen *et al.*, 1990), altered starch composition (Visser *et al.*, 1991), resistance to pathogenic fungi (Keller *et al.*, 1989; Lamb *et al.*, 1992) and bacteria (Anzai *et al.*, 1989), and production of drugs (Mason *et al.*, 1992). Some transgenic crops will be ready for marketing in next few years.

3. DEVELOPMENT OF MARKER GENES IN HIGHER PLANTS

The most important improvements of the Ti plasmid system are based on the development of marker gene or reporter gene systems (Weising *et al.*, 1988). Two of the most popular reporter genes are the β-glucuronidase (GUS) from *E. coli* encoded by the *uidA* locus (Jefferson et al 1987; Jefferson, 1987), the bacterial luciferase (LUX) from a marine bacterium *V. harveyi* encoded by the *luxA* and *luxB* cistrons (Belas *et al.*, 1982; Baldwin *et al.*, 1984; Koncz *et al.*, 1987) and more recently, the *lux A&B* fusion gene, *lux F* (Olsson *et al.*, 1989; Escher *et al.*, 1989)

These genes permit analysis of gene regulation by fusion with regulatory sequences. This method, referred to as gene fusion, is a general approach to study the temporal and spatial regulation of gene expression and to delineate regulatory DNA sequences both in prokaryotic or eukaryotic organisms

(Koncz *et al.*, 1990). The expression of GUS in transgenic plants can be measured accurately using fluorometric assays of very small amounts of plant tissue. Plants expressing GUS are normal, healthy and fertile. Further, GUS is very stable, and tissue extracts continue to show high levels of GUS activity after prolonged storage. Histochemical analysis can also be used to demonstrated the localization of gene activity in cells and tissues of transgenic plants (Jefferson *et al.*, 1987; Jefferson, 1987). Although GUS is convenient to use as a reporter gene in most cases, there are two problems. Native GUS activity has been detected in some species and GUS assays are destructive.

In order to overcome this problem, a novel light-emitting bacterial luciferase reporter gene system (LUX) has been widely used to study the regulation of gene expression in bacteria, yeast and plant cells, plant tissues and intact plants (Koncz *et al.*, 1990; Langridge *et al.*, 1991). It has been known that a number of sensitive assays, such as determination of the concentration of O_2, anaesthetics, antibiotics, and mutagens can be carried out *in vivo* by expression of luciferase in living cells (Hastings and Nealson, 1977). Thus, cloning of luciferase genes provided a tool for novel applications in molecular biology (Baldwin *et al.*, 1984; Cohn *et al.*, 1985).

The luciferase from the marine bioluminescent bacterium *V. harveyi* is a heterodimeric enzyme. The principle underlying the assay procedures is that bacterial luciferase catalyses oxidation of reduced flavin mononucleotide ($FMNH_2$) in the presence of oxygen and the ten carbon aldehyde, decanal, to yield FMN, decanoic acid, water and a photon of light at 490 nm ($FMNH_2 + O_2 + RCHO = FMN + RCOOH + H_2O + 0.1\ h\nu_{490nm}$) (Ziegler and Baldwin, 1981; Kurfurst *et al.*, 1984). Therefore, light can be monitored visually, photographically, or electronically at different sensitivities (Van Dyke, 1985).

For example, light can be captured by X-ray film, a photomultiplier tube or an image intensifier coupled to a video camera for *in vivo* measurements. The expression of bacterial luciferase genes was also studied in plants by dissecting the *lux A* and *lux B* coding sequences from the *lux* operon and fusing them separately to 1' and 2' promoters of mannopine synthase genes. Following transformation, light emission and detection of luciferase subunits showed that assembly of functional luciferase occurred in the cytoplasm of the transformed plant cells (Koncz *et al.*, 1987). Since then, the bacterial luciferase genes have been used successfully to study hormonal regulation (Langridge *et al.*, 1989), analyze promoters (Michel *et al.*, 1993) and carry out T-DNA tagging in higher plants (Jiang *et al.*, 1992; 1993b). All these findings, and further improvements of the *Agrobacterum*/Ti plasmid system provide opportunities for its application to genetic engineering in crop improvement.

4. REGULATION OF PLANT GENE EXPRESSION

Isolation and characterization of sequences involved in regulation of gene expression have been a major focal area of plant molecular biology in recent years. Such analyses have been greatly facilitated by the development of efficient plant transformation systems which allow functional testing of intact and modified gene constructs in transgenic plants to be carried out. These studies have demonstrated that, although plant gene expression is regulated at both transcriptional and post-transcriptional levels, the temporal and spatial regulation of most cloned genes analyzed to date is controlled primarily at the level of transcription (Benfey and Chua, 1989; Kuhlemeier, 1992). As this thesis focuses on the insertional activation of a promoterless bacterial luciferase gene through creation of transcriptional or translational fusions, the properties of transcriptional or translational gene regulation and

their relationships with the level of gene expression are reviewed in the following section.

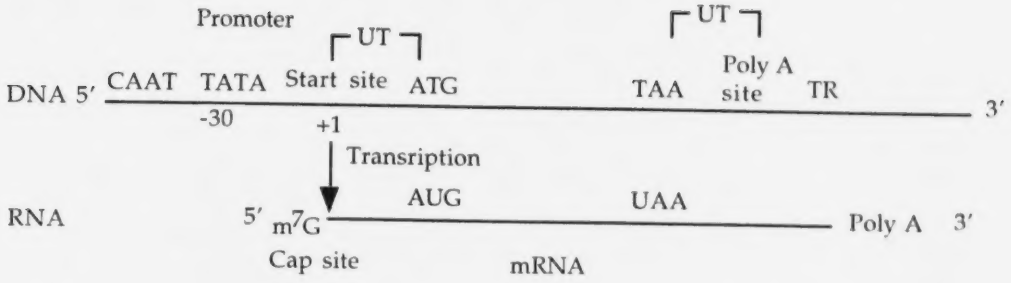
4. A. Transcriptional regulation of gene expression

4. A. 1. *Cis*-acting elements

Transcription is the first stage in gene expression and the principal step at which it is controlled. It has been known that both *cis*-acting elements and *trans*-acting factors are involved in this process (Benfey and Chua, 1989; Kuhlemeier, 1992). The best-characterized of up-stream elements known to be important for gene regulation is the promoter: the region immediately up-stream of each transcription initiation site. A key question in examining the interaction between an RNA polymerase and its promoter is how the protein recognizes a specific promoter sequence. Binding at the promoter is followed rapidly by initiation of transcription at the start site. The start site is defined as the DNA base pair corresponding to the first nucleotide incorporated into RNA. Then RNA polymerase continues along the template until it reaches a terminator sequence. This process results in the production of a single RNA molecule that extends from the promoter to the terminator. This action defines a transcription unit which may include only one or several genes. In higher plants, a transcription unit contains only one cistron (see Fig. II-1).

Two classes of *cis*-acting elements can be distinguished: the "promoter-proper" and the enhancer (Maniatis *et al.*, 1987; Dynan, 1989). Both elements are composed of several recognition sites and function by binding specific proteins (Dynan, 1989). Components of the promoter-proper are located within 100 bp upstream from the transcriptional start site. These include the TATA box, located approximately 30 bp from the start site, which is involved in determining the precise start site of transcription (Elliston and Messing, 1988). Other elements within the promoter-proper modulate levels of

Fig. II-1. Diagram a typical mRNA transcript. The 5′ and 3′ ends of all processed mRNAs contain regions that are not translated into protein. These are called untranslated regions (UTR). A transcription unit for eukaryotic mRNA extends from a transcription start site (Cap site designated as +1) located at the 5′ end of the UTR to a transcription termination region (TR) at the 3′ end. Thus, the nucleotide sequence of the mRNA from the cap site (+1) to the TR should be complementary to the DNA sequence (minus introns). A complex methylated structure (Cap site) is located at the 5′ end of the mRNA and poly A is found at the 3′ end. The Cap consists of a terminal nucleotid, 7-methylguanylate (m7G), in a 5′ - 5′ linkage with the initial nucleotide of the mRNA chain. The Cap site is usually located at a variable location upstream of a ATG translational start codon (AUG in the mRNA). The poly A sequence containing up to several hundreds nucleotides is not encoded by the DNA and is added during the posttranscriptional modification. TAA, TAG or TGA (UAA, UAG or UGA in the mRNA) is a translational stop codon. The TATA box in a plant promoter is located at 30 bp upstream of transcription start site designated as -30.


Promoter
UT
UT
DNA 5′
CAAT
TATA
Start site
ATG
TAA
Poly A site
TR
3′
-30
+1
Transription
RNA
5′
m7G
AUG
UAA
Poly A
3′
Cap site
mRNA

transcription. The best known of these elements is the CAAT box. However, not all plant promoters analyzed contain this element (Elliston and Messing, 1988). Other consensus sequences found within the promoter-proper include the AGGA box, found in the promoter of zein genes (Elliston and Messing, 1988) and the GATA box, found between the TATA and CAAT boxes of several photoregulated gene promoters (Gidoni *et al.*, 1989; Gilmartin *et al.*, 1990). Elements located within the promoter-proper can also confer organ-specific (Stougaard *et al.*, 1987) or light-inducible (Kuhlemeier *et al.*, 1989; Sheen, 1991) regulation upon a constitutive enhancer. Thus, promoter-proper regions from different genes are not functionally equivalent, and cannot always be substituted by corresponding sequences from a different promoter (Castresana *et al.*, 1988).

It is important to note, however, that a promoter alone cannot always drive efficient expression of a gene and in some cases fails to give any detectable expression in transgenic plants (An *et al.*, 1986; Kuhlemeier *et al.*, 1989; Bustos *et al.*, 1991). The additional *cis*-elements required to potentiate transcription are generally referred to as enhancers. Unlike promoter-proper elements, enhancers function in a position- and orientation-independent manner (Maniatis *et al.*, 1987; Dynan, 1989). In some cases, enhancers and promoter-proper elements have been shown to overlap (Benfey *et al.*, 1989; Gilmartin *et al.*, 1990). Although enhancers found in animal gene promoters are able to regulate gene transcription when located 3' of the coding region (Maniatis *et al.*, 1987; Dynan, 1989), most plant enhancers characterized to date have either weak (Fang et al., 1989) or no activity when positioned 3' of a coding region, or have not been tested in this context (Donald and Cashmore, 1990). Therefore, distal regulatory elements found in plants are referred as "enhancer-like" elements (Okamuro and Goldberg, 1989) and plant enhancers

may be functionally more related to upstream activation sites (UASs) found in yeast (Struhl, 1987; Guarente, 1988).

An enhancer typically contains a collection of sites that can bind *trans*-acting factors such as activating or suppressing protein factors. The basic unit capable of forming a binding site for a *trans*-acting factor is referred to as the "core" motif or enhanson (Ondek *et al.*, 1988). Individual core motifs, tandem repeated motifs or two non-identical motifs combine to form the next level of organization, referred to as a domain (Wasylyk, 1988), element (Ondek *et al.*, 1988) or module (Dynan, 1989). Specific combinations of one or several enhancer modules then interact to form an active enhancer. The term enhancer is used for almost any piece of DNA that stimulates transcription. However, it has also become clear that the presence of one or more enhancers sometimes is not sufficient to activate transcription (Dillon and Grosveld, 1993).

4. A. 2. Methods for dissection of *cis*-acting elements

The advantage of gene transfer in higher plants has provided a powerful tool for the understanding of the regulation of gene expression and for the elucidation of basic elements involved in this process. A variety of approaches, including comparison of DNA sequences, mutagenesis followed by functional testing in transgenic plants, and protein binding assays have been used to define and dissect promoters from several plant genes. Based on these studies, it has been proposed that promoters constitute a collection of recognition sites for sequence-specific DNA binding proteins (Kuhlemeier *et al.*, 1987).

Functional analysis of 5' flanking regions has identified DNA sequences that are necessary for gene expression (Maniatis *et al.*, 1987). The principal method used has been to fuse 5' flanking sequences to the protein coding

region of a reporter gene (such as CAT, GUS or LUX) and analyze the expression of the chimeric constructs in transgenic plants or by transient expression assays (electroporation of protoplasts or microprojectile bombardment of plant tissue). The transcriptional activity of these DNA sequences can be inferred from the accumulation of mRNA or protein derived from the chimeric construct. The location of these *cis*-acting sequences can be more precisely defined by deletion and site-directed mutagenesis.

Deletion analyses have been successfully used to delineate promoter regions required for gene activation. For example, a deletion analysis of the 35S promoter of the CaMV showed that the 350 bp upstream region of the TATA box were sufficient for high expression of the reporter gene (Odell *et al.*, 1985). Most recently, a deletion analysis of a promoter region of *rd29A* of *Arabidopsis* in transgenic *Arabidopsis* and tobacco revealed that the -274 to -113 region is essential for the expression of *rd29A* (Yamaguchi-Shinozaki and Shinizaki, 1994). However, deletion analysis alone cannot resolve which specific sequences within promoter are involved in controlling gene transcription. Therefore, other methods such as site-directed mutagenesis or short internal deletions and protein binding assays have been used to correlate the binding of nuclear factors and the ability to activate gene transcription *in vivo*. (Lam *et al.*, 1989; Block *et al.*, 1990; Gilmartin *et al.*, 1990).

Site-directed mutagenesis methods have been developed for introducing mutations into cloned DNA segments, all of which involve the use of enzymes and chemicals that cleave, degrade, or synthesize DNA (Shortle *et al.*, 1981; Itakura *et al.*, 1984; Smith, 1985). Site-directed mutagenesis has been extensively used for many purposes, including analyses

and identification of *cis*-acting elements in plant promoters. Combined with other methods, many motifs in plant promoters that interact with protein factors have been identified by site-directed mutagenesis (Benfey and Chua, 1989; Kuhlemeier, 1992).

There is ample evidence in the literature to demonstrate that enhancer sequences located within plant promoter regions have the ability to bind nuclear factors. The most direct method of identifying such sequences is by binding assays. At present, several techniques are available to specifically analyze protein-DNA interactions, including electrophoretic mobility shift assays (Garner and Revzin, 1981; Fried and Crothers, 1981), DNAse I footprinting (Galas and Shhmitz, 1978) and exonuclease III mapping (Wu, 1985). Gel mobility shift experiments are usually employed to assay the ability of a particular DNA sequence to bind a protein(s), the other techniques can be used alone or in combination to precisely define the nucleotides of the DNA motif that interact with the protein factor. Once a specific DNA sequence has been characterized as a binding site for a particular protein, this motif can be used as a ligand for affinity chromatographic purification of the DNA-binding protein (Kadonaga and Tjian, 1986; Diffley and Stillman, 1986) and as a probe to screen a cDNA expression library (Singh *et al.*, 1988).

Binding assays may be performed *in vitro* or *in vivo*, although results obtained by the two approaches are not always similar. For example, studies in plants have established that sequences which bind protein factors *in vitro* do not always do so *in vivo* (Mckendree *et al.*, 1990). Presumably the presence of a *trans*-acting factor in cell or tissue is not directly correlated to its ability to bind at a recognition site within the promoter of any particular gene. By combining all these methods, Yamaguchi-Shinozaki and Shinizaki (1994)

identified a novel *cis*-acting element in *rd29A* gene of *Arabidopsis* that was involved in responsiveness to drought, low-temperature, or high-salt stress.

In addition, gain-of-function experiments, where a sequence suspected of having regulatory properties is added to a non-active promoter fragment, have been used to test whether a certain motif is required for the initiation of gene transcription. One advantage of this approach is that it can demonstrate whether a motif by itself is sufficient for transcriptional activity. For example, there are several sequence motifs within the enhancer region of the pea *rbc*S promoter, designated GT-1 sites, which bind nuclear factors. Deletions or site-directed mutations within these boxes result in loss of transcriptional activity (Kuhlemeier *et al.*, 1987), indicating a functional role for GT-1 binding. Other gain-of-function experiments were carried out by ligating a tetramer of GT-1 motifs to the promoter-proper of a *rbc*S gene and the CaMV 35S RNA gene. Such constructs could not activate transcription, indicating that GT-1 is necessary, but not sufficient for initiating gene transcription (Gilmartin *et al.*, 1990). Presumably, these GT-1 sequences must interact with other elements within the *rbc*S enhancer. The importance of interactions between GT-1 motifs and other *cis*-acting elements has also been demonstrated by other researchers (Lam and Chua, 1990).

The transcription start site at the 5' end in a protein coding gene can be determined by using the primer extension (Mcknight *et al.*, 1981) or the endonuclease S1 mapping techniques (Berk and Sharp, 1977). In the primer extension technique, an oligonucleotide (primer) which is complementary to an mRNA is prepared and end-labeled. After it is hybridized to the mRNA, the primer can be extended by the enzyme reverse transcriptase until it reaches the first nucleotide of the mRNA. The length of this extension product can then be determined accurately by gel electrophoresis by

comparing with a sequencing ladder derived from the DNA with the same oligonucleotide primer. The location of the TATA box also can be determined. The transcription start site in DNA can also be determined by using the endonuclease S1 mapping techniques (Berk and Sharp, 1977). When the RNA is hybridized with the denatured strands of DNA, it forms an RNA-DNA hybrid with the strand that acted as its template; the other strand of DNA remains unpaired. Treatment with the enzyme S1 nuclease, which specifically degrades single-strand DNA, destroys both the unpaired strand and regions of the template strand beyond the transcription unit. The RNA of the RNA-DNA hybrid can be removed. Then the sequence of the DNA can be determined or its length can be used to locate the position of the RNA on the template. The S1 nuclease mapping techniques can also be used to determine the termination region at the 3' end in a protein coding gene.

4.B. Post-transcriptional regulation of gene expression

In the past, much of the effort to understand gene regulation was devoted to transcriptional mechanisms. It has become increasingly clear that post-transcriptional mechanisms also play important roles in the control of gene expression. Instead of relying solely on transcription, an organism can potentially regulate expression of a gene at multiple levels (Sullivan and Green, 1993). For example, by analysis of accumulation of isocitrate lyase and malate synthase mRNAs in *Brassica napus* L. and run-on transcriptional assays with isolated nuclei, Comai *et al* (1989) indicated that although these two genes were primarily regulated at the transcriptional level, post-transcriptional process also affect the level of mRNA accumulation. Further study on these two genes also suggested that post-transcriptional, translational, and post-translational regulation also differentially affected the

accumulations of mRNAs and proteins in maturing and germinated rapeseeds (Ettinger and Harada, 1990)

Post-transcriptional regulation of gene expression may take place at several different stages, including RNA splicing, processing, transport, and turnover (Okamuro and Goldberg, 1989; Sullivan and Green, 1993). Additionally, regulation at the level of protein translation, post-translational processing and targeting, are by definition, at the post-transcriptional level. Factors involved in plant gene expression at the post-transcriptional level have not been analyzed as intensively as those involved in the transcriptional regulation. This does not imply that post-transcriptional processes are not important for gene expression. In fact, results obtained with cloned genes suggest that post-transcriptional regulation mostly influences the quantitative expression of a gene (Kuhlemeier, 1992; Sullivan and Green, 1993), as some foreign proteins accumulate only to low levels in plants, despite having their genes driven by strong promoters. For example, post-transcriptional regulation is considered to limit the expression of bacterial insecticidal protein (BT-toxin) gene in plants because substantial changes in the transcribed sequences appear to be required for optimal expression (Perlak *et al.*, 1991).

Most studies of post-transcriptional regulation with cloned plant genes have been carried out with gene fusions (transcriptional and translational fusions, see 5. A. for definition). In addition, to simplify measurements of gene activity, the use of defined gene fusions enables researchers to assess the role of various DNA fragments on gene expression by adding and removing these sequences, or modified versions of them. Using gene fusion, several components such as the 5' cap site (Gallie *et al.*, 1991), leader sequences (Futterer *et al.*, 1990; Gallie *et al.*, 1991), initiation codons (Taylor *et al.*, 1987),

introns (Callis *et al.*, 1987; Mascarenhas *et al.*, 1990), 3' untranslated regions (Gallie *et al.*, 1991) and polyadenylation sites (Thornburg *et al.*, 1987; Ingebrecht *et al.*, 1989) have been shown to influence levels of gene expression.

Although the mechanism of mRNA degradation is largely unknown in plants, considerable evidence indicates that mRNA stability contributes to control many plant genes (Sullivan and Green, 1993). For example, it has been reported that DST elements (AUUUA), highly conserved sequence motifs present in the 3' untranslated region (UTR) of a unstable mRNA transcribed from the SAUR (small auxin up RNA) genes in both soybean and *Arabidopsis* can function as instability determinant in plant cells (Newman *et al.*, 1993). The insertion of DST sequences into the 3' UTR, was sufficient to destabilize both the GUS and the globin transcripts in tobacco. By exchanging sequences downstream of the translation start site in different members of the *petunia rbc*S gene family, high to low levels of gene expression were observed (Dean *et al.*, 1989). This result indicates that sequences located 3' of the start site, and probably 3' of the coding region may influence mRNA stability or the relative transcription rate of the *rbc*S genes. Additional evidence for post-transcriptional control that may occur at the level of mRNA stability comes from the finding that several sequences located downstream of transcriptional start sites can affect gene expression, including the 5' UTR of the Em gene of wheat (Marcotte *et al.*, 1989) and a 230 bp sequence encoding the 5' UTR and beginning of the coding region of the Fed-1 gene in pea (Dickey *et al.*, 1992).

In addition to the mRNA stability, translational regulation can be important too. An example of very well documented translational regulation of nuclear gene expression is provided by the *Amaranthus rbcS* genes (Berry

et al., 1988 and 1990). Although light-regulation of *rbcS* genes is mainly controlled at transcription, light appears to affect the stability of certain transcripts. After transferring seedlings from light to dark, mRNA levels for both *rbcS* and the chloroplast-encoded *rbcL* subunits remain unchanged for at least 6 hrs. On the other hand, incorporation of ^{35}S-methionine in the encoded proteins ceases completely within 2 hrs. Subsequent experiments demonstrated that the mRNA remains bound to polysomes, implicating regulation at the level of translation elongation. In contrast, when seedlings were transferred from dark to light, recruitment of *rbcS* mRNA into polysomes was observed, indicating regulation at translation initiation step (Berry *et al.*, 1990).

5. T-DNA-MEDIATED GENE TAGGING

T-DNA-mediated gene tagging is one molecular genetic approach to isolating genes from higher plants (Feldmann, 1991; Walden *et al.*, 1991). It has been successfully used to isolate a number of genes recently (Koncz *et al.*, 1992). In this section, the historical background, the mechanics, strategies and applications of T-DNA-mediated gene tagging are reviewed.

5.A. Historical background of T-DNA-mediated gene tagging

Gene tagging is a technique by which mobile or introduced DNA, with a known sequence, is made to insert into a gene, resulting in a detectable mutation. This serves as a molecular tag for the isolation of the flanking plant DNA. No prior knowledge of the identity of the gene product is required, because selection of insertional mutants is based on visible changes in phenotype . Based on this broad definition, all insertional elements can be used for this purpose, including P elements in *Drosophila* (Bingham *et al.*, 1981), retroviruses in animals (Gridly *et al.*, 1987), insertional elements in

yeast (Roeder *et al.*, 1980), transposons in higher plants (Federoff, 1984) and T-DNAs of *Agrobacterium* (Walden *et al.*, 1991). In this section, discussion is focused on gene tagging with transposons and with T-DNAs of *Agrobacterium*, with emphases on comparing T-DNA-mediated and transposon-mediated gene tagging.

McClintock (1951) discovered the first mobile element in maize over 40 years ago and developed the concept of transposable elements as mobile genetic entities. The mobile elements could move into and back out of genes, altering genetic activity during the process, but these elements were not understood in molecular terms until similar elements were detected in bacteria (for review see Starlinger and Saedler, 1972). The first mobile elements, called insertion sequences (IS elements), were detected through their effect on gene function (Jordan *et al.*, 1968). When IS elements are inserted into the middle of a gene, they inactivate the gene; when they are removed from the gene, the gene regains its activity. For example, *E. coli* loses the ability to grow on galactose after the insertion of an element called IS_2, but the cell regains the ability to grow on galactose upon the removal of IS_2. At least six different IS elements have been found in *E. coli* , and other bacteria also contain them (Starlinger and Saedler, 1972).

Studies of mobile elements in both prokaryotes and eukaryotes have shown that insertion of any mobile element, such as IS elements, transposons, P elements, or retroposons, generally damages the invaded gene (Starlinger and Saedler, 1972; Bingham *et al.*, 1981; Gridly *et al.*, 1987; Walbot, 1992). For example, in yeast, type II alcohol dehydrogenase (ADH_{II}) is an induced enzyme that is synthesized when cells are grown in the presence of alcohol. Glucose represses the synthesis of ADH_{II} even in the presence of alcohol. A Ty element inserted upstream from ADH_{II} gene in yeast results in

a mutation leading to constitutive production of ADH_{II} (Williamson, *et al.*, 1981). In *Drosophila* , eye-color mutations caused by a P element insertion at the *white locus* have been reported (Levis *et al.*, 1984). These and other studies in insertional mutagenesis with mobile elements have led to isolate genes from both prokaryotes and eukaryotes. Transposons, for example, have been used to isolate a variety of genes in bacteria (Kleckner, 1981), *Caenorhabditis elegans* (Greenwald, 1985), *Drosophila* (Cooley *et al.*, 1988), and mouse (Gridley *et al.*, 1987).

In higher plants, based on the type of insertional elements used, two different gene tagging systems have been developed and extensively studied. These are transposon-mediated and T-DNA-mediated gene tagging (Walbot, 1992). Transposons, best known through the work of McClintock (1951), are well-defined DNA segments which can move from one location to another in an organism's genome. As a consequence of the movement of transposons around the genome, a high frequency of spontaneous mutations are caused by their insertion within genes. Individuals that have inherited the mutation are identified by screening for the appropriate phenotype. All native transposons in higher plants encode functions required for their transposition (transposases) which recognize specific sequences at the ends of the transposons (Federoff, 1984). Since they were isolated from maize (*Zea mays*) (Fedoroff *et al.*, 1983) and snapdragon (*Antirrhinum majus*) (Sommer *et al.*, 1985), transposons have been successfully used to isolate a variety of genes in these two species that possess well-characterized endogenous transposons (for review see Walbot, 1992). However, in most species of higher plants, lines with active transposons are either not available or have not been identified (Gierl and Saedler, 1992). To overcome this limitation and allow transposon tagging in any species, transposons from maize and

snapdragon have been introduced by transformation into a variety of heterologous hosts, including tobacco (Baker, *et al* ., 1986), tomato (Yoder, 1988), carrot (Van Sluys *et al.*, 1987), *Arabidopsis* (Van Sluys *et al.*, 1987; Bancroft *et al.*, 1993), rice (Murai *et al.*, 1991) and petunia (Chuck *et al.*, 1993). Although it has been reported that transposons were mobile in some heterologous hosts (Walbot, 1992), only two recent reports have shown that genes were tagged by transposons in *Arabidopsis* (Bancroft *et al.*, 1993) and petunia (Chuck *et al.*, 1993). Therefore, the development of an efficient transposon tagging system in heterologous hosts is continuing to be an important objective in this area.

Alternatively, T-DNAs of *Agrobacterium* can be used for gene tagging in higher plants (Andre *et al.*, 1986; Terri *et al.*, 1986; Walden, 1991). This molecular genetic approach was directly developed from the interaction of *Agrobacterium* research with plant molecular biology (Schell, 1987; Koncz *et al.*, 1992). As was discussed early, many of the properties of the T-DNA showed that T-DNA can be a suitable candidate employed for gene tagging. T-DNA used for gene tagging is based on the fact that it is stably integrated into the plant nuclear genome, hence its insertion into genes may cause mutations (Zambryski *et al.*, 1989). Therefore, the T-DNA border-based stable integration of a known foreign DNA sequence into a plant genome may serve both as a mutagen and a molecular tag (Andre *et al.*, 1986; Koncz *et al.*, 1989; Feldmann *et al.*, 1989; Walbot, 1992). Like transposons, *Agrobacterium*-mediated T-DNA integration may result in detectable mutations through the insertional inactivation of a functional gene. These inserts do not move once inserted. Thus, unlike transposon-mediated gene tagging which depends on the relocation of transposons for creating mutations, T-DNA-mediated gene tagging relies on generating more transgenic individuals for mutant

selection. Unlike transposons, T-DNA does not encode any functions required for its transfer and integration into the plant genome. A reflection of the differences on integration and structure between T-DNA and transposons is apparent when they are used for gene tagging. For example, any DNA introduced between the T-DNA borders will transfer into plant cells offering more applicable strategies in T-DNA tagging. This is not the case for transposons, where integrity of internal sequences is generally important for transposition. However, this limitation in transposons has been overcome the development of a two component system (Coupland *et al.*, 1988). Another advantage is that the copy number of T-DNA inserts in a genome is small and the integrated T-DNA is stable, whereas, the copy number of transposons is usually high and unstable. Once inserted, the T-DNA will not excise (Zambryski, 1988). However, the lack of movement of T-DNA can be perceive as an advantage and also a disadvantage. It is an advantageous for gene cloning purpose, but makes it impossible to generate multiple alleles by reversion. Therefore, it is much easier to identify and clone flanking plant DNAs with T-DNA than with transposons. On the other hand, gene tagging with a transposon could reduce the time and effort needed to transform and regenerate plants, because subsequent integration of an transposon at new chromosomal locations will create new mutation.

The possibility of T-DNA-mediated gene tagging was first demonstrated by Andre *et al* (1986) and Terri *et al* (1986) and developed by Koncz *et al* (1989) and Feldmann *et al* (1989). A promoterless kanamycin resistant gene was linked to the right border of T-DNA and transformed into *Nicotiana tabacum* and haploid *N. plumbaginifolia* plants. Calli expressing T-DNA-mediated gene fusions were identified by their ability to grow in the presence of kanamycin. Subsequently, NPT II assays were used to select regenerated

individuals with organ-specific gene expression. Selection for kanamycin resistance in both species resulted in transcriptional activation of the promoterless reporter gene (Andre *et al.*, 1986; Terri *et al.*, 1986). In the same experiment, plants transformed by a vector containing a reporter gene lacking the ATG initiation codon synthesized diverse kanamycin phosphotransferase fusion proteins. These data indicated that T-DNA integration can occur in protein coding regions of the nuclear genome creating active transcriptional or translational plant gene-reporter gene fusions. Some morphological altered *N. plumbaginifolia* plants were obtained (Andre *et al.*, 1986; Terri *et al.*, 1986). These early studies suggest that T-DNA tagging may become a suitable complement in all plants for which *Agrobacterium*-mediated transformation, tissue culture and regeneration methods are established.

To identify T-DNA insertions in functional plant genes, an *in vivo* gene fusion technique was developed (Andre *et al.*, 1986; Terri *et al.*, 1986) and improved (Koncz *et al.*, 1989; Kertbundit *et al.*, 1991). Simply stated, a gene fusion is a chimeric DNA construct containing sequences from more than one gene. Gene fusions may be created *in vitro*, using standard recombinant DNA technology, or *in vivo* by insertions of mobile genetic elements within intact genes. Most of the studies with gene fusions have exploited constructs created *in vitro* that contain the coding region from a reporter gene or selectable marker. Generally speaking, two types of gene fusions are recognized. Transcriptional fusions contain the coding region of the reporter gene only, while translational fusions encode chimeric proteins containing segments from both the reporter gene and the gene of interest resulting in "in frame" translation of both genes (Jefferson, 1987).

The major advantage of the above approach is that a direct selection of promoterless kanamycin and hygromycin resistant genes for the presence of

T-DNA-mediated gene fusions greatly facilitates identification of desired transformants. This becomes especially important when gene fusions are generated at low frequencies. There are, however, several disadvantages with this approach. If expression of gene fusions is selected during the tissue culture stage, it is possible to lose individuals with expression of gene-fusion at subsequent stages of plant development. Second, it has been reported that the direct selection of promoterless kanamycin and hygromycin resistant genes for transformants resulted in a high copy number of T-DNA insertions (Koncz *et al.*, 1992). Finally, when promoterless kanamycin or hygromycin resistant genes are place into the middle of a T-DNA insert, direct selection may also result in a high frequency of truncated and aberrant T-DNAs (Herman *et al.* 1990).

In order to avoid direct selection for expression of gene fusions, a major improvement of T-DNA tagging vector systems was developed by Koncz *et al.* (1989). T-DNA vectors carrying a promoterless reporter gene (*nptII* gene) were linked to the right border of the T-DNA and a separate selectable marker (hygromycin-resistant gene) for recovery of transgenic calli and plants were employed. A ColE1 replicon and a bacterial selectable marker gene were included in these vectors to facilitate the recovery of T-DNA inserts by plasmid rescue (Koncz *et al.*, 1989). Using this strategy, a large number of transgenic *Arabidopsis* and tobacco plants were generated and a root-hair specific promoter (Koncz *et al.*, 1989) and a pale mutant from *Arabidopsis* (Koncz *et al.*, 1990) were isolated. However, quantification of NPTII is laborious and no histochemical technique for its detection is available. A more attractive alternative for the isolation of random gene fusion was developed by using a promoterless GUS gene to construct T-DNA tagging vectors (Kertbunit *et al.*, 1991). In this study, the promoterless GUS gene was

located 4 bp from the right T-DNA border to generate transcriptional and translational fusion upon T-DNA integration (Kertbunit *et al.*, 1991).

An other improvement of T-DNA tagging vector systems is the use of an enhancer detector, a marker gene linked to a minimal promoter, which has been successfully used in transgenic mouse (Allen *et al.*, 1988), *Drosophila* (Wilson *et al.*, 1989) and plants (Goldsbrough and Bevan, 1991; Topping *et al.*, 1991). A truncated CaMV 35S promoter (-90) fused with a GUS gene placed the left and right T-DNA borders, led to a high frequency of transgenic tomato (Goldsbrough and Bevan, 1991) and tobacco (Topping *et al.*, 1991) plants that displayed differential patterns of GUS expression.

Most protocols used for generation of T-DNA-tagged plant populations involve extensive periods of tissue culture. It has been reported that tissue culture processes induce a high frequency of somaclonal variants in the plant population (Feldmann and Marks, 1986). The seed transformation method in *Arabidopsis* developed by Feldmann and Marks (1987) circumvents potential problems related to tissue culture. Using this method, more than 1000 putative insertional mutants were produced (Feldmann, 1991). Many of the T-DNA insertional mutants turned out to be allelic to known genes, and some of these genes were isolated using the T-DNA as a tag, including *Gl1* (glabrous) (Marks and Feldmann, 1989), *AG* (agamous) (Yanofsky *et al.*, 1990), *cop* 1 (Deng *et al.*, 1992) and *CTR1* (Kieber *et al.*, 1993).

5. B. Mechanics and strategies of T-DNA-mediated gene tagging

5. B. 1. Structure and organization of the integrated T-DNA

Currently little is known about the fate of T-DNA during transformation. Most of the information available is obtained from the analysis of the end product of the integrated T-DNA in transgenic plants. Since sequences contained within the T-DNA are usually well-characterized,

it is relatively straightforward to use these as hybridization probes to study the structure of the integrated T-DNA. These analyses provide information on copy number and rearrangement of the integrated T-DNA. In addition, sequence analysis and comparison of T-DNA-plant DNA junctions provide further information about the T-DNA integration.

Southern blot analyses and sequence comparisons of the T-DNA-plant DNA junctions of transgenic plants have revealed that T-DNA integration into plant chromosomes is a very precise event. The endpoints of many integrated T-DNA molecules are located within or less than 10 bp away from the right border of the T-DNA (Jorgensen *et al.*, 1987; Gheysen *et al.*, 1991; Mayerhofer *et al.*, 1991). In comparison, integration at the left border of the T-DNA does not appear to be as precise. The Vir D2 protein was found covalently attached to the 5' end of the right T-DNA terminus in *Agrobacterium* (Herrera-Estrella *et al.*, 1988; Young and Nester, 1988). These results imply that the right border of the T-DNA may be protected during transfer.

At the chromosomal level, *in situ* hybridization and mapping using genetic or RFLP markers have established that no preferential target for T-DNA integration could be observed (Zambryski *et al.*, 1989). However, at the gene level, recent results from T-DNA tagging experiments based on the frequency of insertional activation (Koncz *et al.*, 1989; Topping *et al.*, 1991), indicated that the T-DNA integrates preferentially into transcribed regions of the plant genome. T-DNA insertion units usually consist of one to several copies of T-DNA. In cases of multiple copies of the T-DNA, independent locations of several inserts and tandem arrangements of T-DNA (which consist of direct and inverted repeats) were observed (Jorgensen *et al.*, 1987; Zambryski, 1988; Zambryski *et al.*, 1989).

Molecular analysis of over 200 transformants revealed that the selection method can influence the copy number of the T-DNA (Koncz *et al.*, 1992). Selective probing of genomic DNAs isolated from these plants with promoterless selectable marker genes (such as kanamycin or hygromycin) from the right border of the T-DNA indicated that on average 5 to 20 T-DNA copies were inserted. In contrast, control plants transformed by promoter-driven selectable marker genes carried intact T-DNA at an average copy number of 1.5 (Koncz *et al.*, 1992). Similar results were also obtained with a promoterless marker gene placed into the middle of a T-DNA insert (Herman *et al.*, 1990). These observations suggest that direct selection for T-DNA integration into actively transcribed plant genes leads to amplification of promoterless marker genes and integration of truncated, aberrant T-DNAs.

5. B. 2. Models of T-DNA integration

Based on comparisons of sequences in junctions of T-DNA to plant DNA with sequences of plant DNA before and after the T-DNA integration, models of T-DNA integration were proposed (Gheysen *et al.*, 1991; Mayerhofer *et al.*, 1991). These models suggest that the protein-T-strand complex invades the target region of the genome by utilizing nicks or gaps in the double-stranded DNA introduced by DNA replication, recombination, or repair. It is still questionable whether any of the *Agrobacterium*-encoded *vir* gene products such as Vir D2 and Vir E2 are involved in this process, although it has been proposed that these proteins may be involved in recognition and nicking of plant DNA, mediating the formation of precise junctions at the right border of the T-DNA, and unwinding of the target DNA (Mayerhofer *et al.*, 1991). Based on these models, any sequence close to the borders of the T-DNA has the potential to anneal with plant DNA resulted in the formation of a heteroduplex. The small size of the target sites indicated

that the borders of the T-DNA are located close to one another during this process. The data from these studies also demonstrate that the process of T-DNA integration does not create large rearrangements of plant DNA. Some small (13-73 bp) deletions of target site DNA and "filler" DNA were often found in the T-DNA-plant DNA junctions. "Filler" DNA are short stretches of repeated sequences, they are different from the plant target site and the T-DNA borders, but usually identical to sequences found in the nearby flanking DNA. The fact that there is little or no homology between the insertion sites suggests that the T-DNA integration is mediated by illegitimate recombination (Gheysen *et al.*, 1991; Mayerhofer *et al.*, 1991).

5. B. 3. Strategies: T-DNA insertional vectors

Generally, all T-DNA based vectors can be used for T-DNA tagging because mutant phenotypes created by T-DNA integration will be segregated in the second generation of the primary transgenic plant. The use of non-specialized T-DNA transformation vectors to generate a population of insertion mutants has been demonstrated (Feldmann *et al.*, 1989; Feldmann, 1991).

In the last few years, a variety of T-DNA tagging vectors have been constructed. These vectors are called specialized T-DNA tagging vectors. A so called 'passive vector', contains a marker gene either lacking a promoter sequence or containing a minimal promoter, i.e missing sequences necessary to produce normal developmentally regulated gene expression, but sufficient to initiate transcription. The rationale is that the following flanking plant sequences will be able to direct expression of the marker gene (Walden *et al.*, 1990 and 1991). The *nptII* gene (Andre *et al.*, 1986; Teeri *et al.*, 1986), hygromycin gene (*hpt*) (Koncz *et al.*, 1989) and GUS gene (Kertbundit *et al.*, 1991; Topping *et al.*, 1991) have all been used to construct passive vectors.

Other types of specialized T-DNA tagging vectors are called 'active vectors'. They contain transcriptional enhancer sequences which can enhance gene expression over a long distance (Walden *et al.*, 1991). The transcriptional enhancer derived from the cauliflower mosaic virus 35S RNA promoter placed close to the right T-DNA border is an example of the 'active' vector and can enhance gene expression (Walden *et al.*, 1991). This strategy allows the production of dominant, overexpression mutants. Using this method, a plant gene that promotes growth and division of protoplasts *in vitro* in the absence of exogenously applied auxins has been isolated (Hayashi *et al.*, 1992).

5. C. Regulation of transgene expression

Transcriptional and post-transcriptional control of gene expressions have already been discussed in an earlier section. Although the spatial and temporal expression of transgenes usually parallel those of the original gene, levels of expression often vary considerably among independent transformants with identical T-DNA constructs (Hobbs *et al.*, 1990; Peach and Velten, 1991). Altered patterns of gene expression in transgenic plants compared with those of original genes have also been reported (De Almeida *et al.*, 1989; Barnes, 1990). As the transgene expression is a major focal point in this thesis, factors responsible for influencing transgene expression are reviewed in this section.

5. C. 1. Copy number of T-DNA insert and transgene expression

Several factors have been proposed to be responsible for the variability of gene expression including the copy number of T-DNA insert and position effect (Hobbs *et al.*, 1990 and 1993). Surprisingly, most data up to date have shown that there is no positive correlation between levels of gene expression and T-DNA copy number. On the contrary, the presence of multiple copies of

T-DNA inserts, arranged in tandem repeats, appears to have a negative effect on gene expression (Jones *et al.*, 1987; Jorgensen *et al.*, 1987). A negative correlation between levels of gene expression of a CaMV 35S promoter-GUS construct fusion and T-DNA copy number was reported (Hobbs *et al.*, 1990). Their data indicated that even plants with two T-DNA copies do not express GUS to levels comparable of plants containing single inserts. A negative effect of multiple T-DNA copies was also demonstrated by analysis of GUS expression in the transformed progeny. When two independent T-DNA copies segregated away from each other, higher levels of GUS were detected.

The reasons why multiple copies of T-DNA insert have negative influences on transgene expression are largely unknown. Based on the observation that some sequential transformation of plants with identical or different T-DNAs may lead to the inactivation of genes on the original T-DNA, co-supression (Jorgensen, 1990), or *trans*-inactivation (Jorgensen, 1991) was suggested. However, the co-suppression or *trans*-inactivation hypothesis still cannot explain why only some transformants were suppressed.

5. C. 2. Position effect

Most genes whose expression has been studied in transgenic plants are expressed in generally appropriate patterns with respect to cell and organ specificity, developmental timing, and to environmental cues (Benfey and Chua, 1989; Kuhlemeier, 1992). However, the level of expression can vary over an extremely wide range (Dean *et al.*, 1988; Hobbs *et al.*, 1990 and 1993). Variation in transgene expression is frequently attributed to corresponding variation in the transcriptional potential of different chromosomal insertion sites (Allen *et al.*, 1993). The phenomenon whereby the presence of other regulatory sequences in the host genome at or near the site of T-DNA integration influences the level and specificity of expression of the transgene

is called the position effect (Dilton and Grosveld, 1993). Besides endogenous promoters and enhancers, other regulatory sequences, including the locus control region (LCR) (Grosveld *et al.*, 1987), matrix attachment regions (MAR) (Dillon and Grosveld, 1993) and chromosomal scaffold attachment regions (SAR) (Gasser and Laemmli, 1986), also influence transgene expression.

Position effects are common in transgene research. Prominent among the many factors that may affect transgene expression at different insertion sites of chromosome is the structure of chromatin. For example, the transcriptional potential of large regions is controlled by chromatin fibre, thus permitting access to the DNA by RNA polymerase and transcriptional regulatory proteins (Goodman, 1988; Cook, 1989). Scaffold attachment regions (SAR) function as domain boundaries and play an important role in regulating transgene expression. It has been reported that the SAR could reduce position effect variation *in vivo* when included on both sides of globin gene construct (Grosveld *et al.*, 1987). Most recently, the SAR has been shown to reduce position effect in transgenic plant cells when placed on both sides of GUS gene (Breyne *et al.*, 1992; Allen *et al.*, 1993).

One method that has been used to compensate for position effects is to add a reference gene within the T-DNA. By comparing the expression of the test gene with that of the reference gene, variability among different transformants may be significantly reduced (Kuhlemeier *et al.*, 1987; Fang *et al.*, 1989). However, in several cases, expression of linked transgenes was shown to vary independently from one another (Gidoni *et al.*, 1988; Peach and Velten, 1991). In one study, the ratio of CAT to GUS activity, when fused with the mannopine synthase promoter (mas) P1 and P2, varied from 0.05 to 49 (Peach and Velten, 1991). If this variability is a result of a position effect,

the influences of flanking regulatory sequences must vary considerably over short distances, as the mas promoters were tightly linked to each other.

Transgene expression may be influenced not only by its position within the plant genome, but also by its position within the T-DNA (Gidoni *et al.*, 1988). The expression of two photosynthetic petunia genes (*Cab22R* and *Cab22L*) was controlled at the level of transcription in a tissue-specific and light-regulated fashion, but was influenced by the site of insertion into the T-DNA vector. When the two transcription units were inserted near the left border of the T-DNA, the expression level varied coordinately among independent transgenic plants and remained completely light inducible and tissue specific. In contrast, the expression level and transcription properties of the same gene construct was significantly disrupted, when it was inserted near the right border of the T-DNA (Gidoni *et al.*, 1988).

Transgene expression may be influenced by the developmental stages, position in the plant and physiology. One study showed that expression of a nos promoter-CAT fusion varied almost 300 fold between leaf tissue harvested from the bottom and top of tobacco shoots (An *et al.*, 1988). The expression of a gene fusion between the β- 1,3-glucanase gene promoter and GUS was found only in older leaves of transgenic tobacco (Castresana *et al.*, 1990).

5. D. Phenotype and genetic analysis of T-DNA inheritance

5. D. 1. Predicted phenotypes and genotypes of primary transformants

Once the T-DNA has been stably integrated within the plant genome, heterozygous primary transformants which contain one or more inserts will be generated. In most cases, as there is no counterpart in the plant genome, the T-DNA tags are dominant (Koncz *et al.*, 1992). Therefore, the primary transformant will express dominant marker genes carried by the T-DNA.

Subsequently, marker genes such as kanamycin and hygromycin which were originally used to select for transformants are easily used to study the segregation of the T-DNA insert. For segregation experiments, seeds harvested from self-pollinated primary transformants are plated on a medium containing kanamycin or hygromycin at the room temperature. In this environment, most seeds will germinate, including those with or without the T-DNA insert(s). Only those containing T-DNA insert(s), however, will develop into normal seedlings, with exception of these homozygous embryo-defective mutants. Approximately 200 seeds from each primary transformant are required. Several hundred more seeds will need to be examined in these non-segregating lines to distinguish between multiple linked or unlinked inserts and homozygosity (Feldmann, 1991).

In a primary transformant (T1) (equivalent to F1), contained one T-DNA insert that carried a kanamycin resistant gene, progeny of the selfed primary transformant would segregate in a Mendelian fashion at T2 (equivalent to F2) under the selection of kanamycin. The ratios of the segregation will be a 3 : 1 (Kan^{RR} : Kan^{rr}) phenotypically and an 1 : 2 : 1 (Kan^{RR} : Kan^{Rr} : Kan^{rr}) genotypically. If a primary transformant contains two unlinked T-DNA, progeny of the selfed primary transformant would still segregate in a Mendelian fashion. The ratios of the segregation would be a 15 : 1 ($Kan^{R_1\text{-}R_2\text{-}}$: $Kan^{r_1r_1r_2r_2}$) phenotypically and an 9 : 3 : 3 : 1 ($Kan^{R_1\text{-}R_2\text{-}}$: $Kan^{R_1\text{-}r_2r_2}$: $Kan^{r_1r_1R_2\text{-}}$: $Kan^{r_1r_1r_2r_2}$) genotypically.

5. D. 2. T-DNA insertional mutagenesis in diploid and amphidiploid species

The efficiency of the T-DNA insertional mutagenesis is affected by many factors, including plant species, genome size, and methods of transformation, regeneration and screening for mutant transformants. Most data available to date on the T-DNA insertional mutagenesis have been generated in two

species, *Arabidopsis thaliana* and *Nicotiana tabacum* (Koncz *et al.*, 1992). Their genomes vary in size and organization and the methods employed to obtain insertional mutants are different. This section focuses on these two species, and reviews the current literature on how the basic genetics of these two species affects the efficiency of mutant selection.

The DNA content of these two species varies considerably. *Arabidopsis thaliana* has the smallest genome characterized to date in higher plants (Meyerowitz, 1987). At 7×10^4 kb, it is only 5 times larger than that of baker's yeast (Leutwiler *et al.*, 1984). On the other hand, the genome size of *Nicotiana tabacum* is 2.4×10^6 kb per haploid genome. This is approximately 25 times larger than that of *Arabidopsis* and is similar to the size of the haploid human genome (2.0×10^6 kb) (Okamuro and Goldberg, 1985). In spite of the difference in total genomic content and percentage of repeated sequences, Koncz *et al* (1989) reported that the probability of generating a T-DNA-mediated gene fusion in *Arabidopsis* and *Nicotiana* was very high and similar, 28.5% vs. 23.5% for transcriptional fusions and 30% vs. 24.4% for translational fusions. Based on their results, these authors suggest that the T-DNA integrated preferentially into transcribed regions of the plant genome (Koncz *et al.*, 1989).

It should be also noted that there was no difference between a diploid species (such as *Arabidopsis thaliana*) and amphidiploid species (such as *Nicotiana tabacum*) in the segregation ratios of T-DNA marker genes. A primary transformant from both species will follow the same segregation principal as discussed above. However, the difference between these two species is obvious when the focus is on screening for insertional mutants. Unlike diploid *Arabidopsis*, the genome of amphidiploid tobacco consists of two complete sets of chromosomes resulting from the interspecific

hybridization of two diploid species, *Nicotiana sylvestris* and *Nicotiana tomentosiformis* (Okamuro and Goldberg, 1985). Theoretically it should be almost impossible to obtain insertional mutants from tobacco, because most genes in this species contained their complementary sets. With the exception of gain-of-function type mutation such as the introduction of reporter genes, the lost function resulting from mutation of a gene may be compensated by its partner. This is a major reasons why most of the early T-DNA tagging experiments failed to isolate insertional mutants from this species (Koncz *et al.*, 1992).

It has been reported that most transformants tested will segregate in a Mendelian fashion and that each transformant will segregate for one or more inserts (Feldmann, 1991; Topping *et al.*, 1991). By testing more than 800 transformed lines of *Arabidopsis,* Feldmann, (1991) demonstrated that the average number of T-DNA inserts per transformant was 1.4 and nearly all of the transformants were heterozygous for one or more inserts. With the exception of one study, a certain fraction of the transformants contained insertions that segregated in a non-Mendelian fashion, including those plants which failed to segregate for a homozygous T-DNA insertion and those plants which segregated for the presence of the T-DNA at low frequency (Deroles and Gardner, 1988). In a recent study, plants failing to segregate in a Mendelian fashion for hygromycin resistance were shown to segregate according to Mendelian principals when the progeny was analyzed by Southern hybridization (Scheid *et al.*, 1991). These results may be explained, at least in part, by inactivation of T-DNA-encoded marker genes.

6. APPLICATIONS OF T-DNA TAGGING IN HIGHER PLANTS

A large number of transgenic plants containing tagged genes have been obtained by *Agrobacterium*-mediated transformation of *Arabidopsis* and tobacco (Koncz *et al.*, 1989; Feldmann, 1991). Developmentally regulated promoters and genes from higher plants have been isolated by T-DNA tagging or T-DNA insertional mutagenesis (Walbot, 1992). A root-specific promoter from *Arabidopsis* has been isolated by using a promoterless *aph(3')-II* reporter gene-based insertional vector (Koncz *et al.*, 1989; 1990). A DNA fragment mediating phloem-specific gene expression in *Arabidopsis* has been isolated by T-DNA tagging with a promoterless GUS gene-based insertional vector (Kertbundit *et al.*, 1991). In the last few years, a number of plant genes isolated by T-DNA tagging, have been studied including 'master' genes such as the *AG* gene in *Arabidopsis* that encodes a transcription factor responsible for flower development (Yanofsky *et al.*, 1990), the *cop* 1 gene that encodes a protein containing both a motif of G protein and a DNA-binding domain (Deng *et al.*, 1992) and a plant gene that, upon overexpression, is able to promote growth and division of protoplasts *in vitro* in absence of exogenously applied auxins (Hayashi *et al.*, 1992).

The leaf disc transformation method (Horsch *et al.*, 1985) and seed transformation method (Feldmann and Marks, 1987; Feldmann, 1991) have been used successfully for creation of a large number of insertional mutant plants. Leaf discs are cocultivated with *A. tumefaciens* containing the appropriate T-DNA tagging vector, followed by culturing the explants on medium containing an antibiotic to prevent further growth of the bacteria while allowing the induction of callus or shoots and roots formation. Transformant selection can be applied at the initiate stage of callus or shoot formation. Using this method a population of transgenic plants has been

generated in both *Arabidopsis* and tobacco (Koncz *et al.*, 1989; Kertbundit et al., 1991; Jiang *et al.*, 1992; Lindsey *et al.*, 1993).

An alternative strategy for creating T-DNA tagged plant populations is seed transformation (Feldmann and Marks, 1987; Feldmann, 1991). In *Arabidopsis*, imbibing seeds are incubated with *A. tumefaciens* and resulting plants are self pollinated. The progeny is germinated in the presence of antibiotics to select for expression of the T-DNA encoded selectable marker gene in transgenic individuals.

Previous studies have shown that a high frequency of T-DNA tagged individuals can be obtained in both *Arabidopsis* (Koncz *et al.*, 1989; Kertbundit *et al.*, 1991) and tobacco (Koncz *et al.*, 1989; Kertbundit *et al.*, 1991; Topping *et al.*, 1991). Depending on the type of T-DNA tagging vector, screening can be based on the mutant phenotype (Feldmann *et al.*, 1989; Feldmann, 1991) and spatial or temporal expression of marker genes in transgenic plants (Kertbundit *et al.*, 1991; Jiang *et al.*, 1992). The ch-42 gene was isolated by mutant phenotype screening of 450 T-DNA-transformed *Arabidopsis* plants (Koncz *et al.*, 1990). Similarly, more than 1000 putative mutants were obtained by screening 8000 *Arabidopsis* transformants for mutant phenotypes generated through seed transformation (Feldmann, 1991).

Identification of insertional mutants through marker gene expression has been reported recently (Koncz *et al.*, 1989; Kertbundit *et al.*, 1991; Topping *et al.*, 1991). A T-DNA vector with a promoterless aminoglycoside phosphotransferase II (*aph(3')-II*) reporter gene at one end was designed to study the frequency of T-DNA tagging both in *Arabidopsis* and in tobacco (Koncz *et al.*, 1989). A minimal promoter and promoterless *uid A* (*gusA*) reporter gene, encoding the enzyme β-glucuronidase (GUS) (Jefferson *et al.*, 1987), was used to isolate developmentally regulated promoter and enhancer

elements (Kertbundit *et al.*, 1991). The GUS reporter gene, fused to a truncated (-90 bp) CaMV35S promoter, was described for identification of tissue-specific promoters in transgenic tobacco plants (Topping *et al.*, 1991). In general, these reporter genes are sensitive and convenient to use, but measurement of enzymic activity requires destructive essay procedures.

Based on the mutant phenotype or spatial or temporary expression of a marker gene in transgenic plants, the tagged individual of interest can be studied in detail. The flanking DNA sequence of interest can be isolated by conventional molecular biology techniques and recent developments using plasmid rescue (Koncz *et al.*, 1989). In all these cases, careful Southern hybridization analysis (Southern, 1975) is required to measure the length of the flanking plant DNA sequence. Cloning of the T-DNA tagged plant DNA sequences can be carried out by screening partial or complete genomic libraries derived from the mutant individual. If the T-DNA itself contains an origin of replication and antibiotic resistance gene in *E. coli*, the flanking plant DNA sequence can be directly rescued in *E. coli* after the cleavage of the genomic DNA with an enzyme that cuts outside the T-DNA. This is followed by religation and bacterial transformation (Koncz and Schell, 1986; Koncz *et al.*, 1989). Using inverse PCR, T-DNA tagged flanking plant DNA sequences can be isolated directly (Jiang et al., 1992).

The flanking plant DNA sequences may be a gene or its regulatory sequence. If the selection of individual mutants is based on the expression of a marker gene, the flanking plant DNA sequence may be involved in the regulation of gene expression. This method is called 'promoter tagging' or 'promoter trapping' (Kertbundit *et al.*, 1991; Lindsey *et al.*, 1993). In this case, the putative promoter sequence can be confirmed by DNA sequence analysis and retransformation experiments and can be used as a probe to screen a

genomic library for the isolation of downstream coding region. On the other hand, if the flanking plant DNA sequence itself is a part of a gene sequence, the above method can also be used to isolate the complete gene and its upstream regulatory sequence (Walden *et al.*, 1991; Koncz *et al.*, 1992).

7. GENETIC CONTROL OF FLOWER DEVELOPMENT

Flowering plants are the most highly evolved and complex organisms within the plant kingdom. The flower of angiosperms is the reproductive organ and consists of sepals, petals, stamens and carpels, which are usually organized in four concentric whorls (Coen and Meyerowitz, 1991). In order to make such an arrangement, the cells of the organ primordia in each of the whorls must spatially deliver the correct developmental program. Recently, genetic analysis of *Arabidopsis* and *Antirrhinum majus* (snapdragon) mutants with abnormal flowers has produced fruitful information about flower development (Schwarz-Sommer *et al.*, 1990; Coen and Meyerowitz, 1991; Meyerowitz *et al*,. 1991). The genes defined by these mutations have been isolated including the homeotic genes *AGMOUS* (*AG*), *APETALA2* (*AP2*), *APETALA3* (*AP3*), and *PISTILLATA* (*PI*). Plants with mutation(s) in one or more these genes produce flowers with altered morphology. In these mutant flowers, organs in some whorls take the form of those normally found in other whorls. The floral homeotic genes fall into three distinct classes, designated A, B, and C. The class A (*AP2*) mutation leads to first whorl organs developing as carpels rather than sepals and second whorl organs developing as stamens rather than petals. The class B (*AP3* or *PI*) mutations result in second whorl sepals and third whorl carpels. The class C (*AG*) mutations convert stamens into petals. These genetic studies of the flower mutations have led to a model explaining the roles of floral homeotic

genes in determining floral organ identity (Coen and Meyerowitz, 1991). The model proposes that three gene products function in three overlapping fields (A, B, and C), with each field occupying two adjacent whorls. Thus, field A consists of whorls 1 and 2, field B consists of whorls 2 and 3, and field C consists of whorls 3 and 4. The *AP2*, *AP3/PI*, and *AG* genes are proposed to function in fields A, B, and C, respectively, resulting in a unique pattern of expression in each of the four whorls (Bowman *et al.*, 1991).

The cloning, expression and function studies of these homeotic genes have shown results similar to the predicted model. Sequence analysis of the class B homeotic genes including *Deficiens-A* (*DEFA*) (Sommer et al., 1990), *Globosa* (*GLO*) (Trobner *et al.*, 1992), *AP3* (Jack *et al.*, 1992) and the class C homeotic gene, *AG*, reveals that they contain a putative DNA-binding domain, designated the MADS (MCM1, AG, DEFA, and SRF) box (Schwarz-Sommer *et al.*, 1990). The MADS box is similar to the DNA-binding domain of the yeast transcription factor MCM1 (Dubois *et al.*, 1987; Christ and Tye, 1991) and the human transcription factor SRF (Norman *et al.*, 1988), therefore, these studies suggest that the classes B and C homeotic genes may encode plant transcription factors which regulate the genes determining floral organ identity.

BIBLIOGRAPHY

Abel, P.P., Nelson, R.S., De, B., Hoffmann, N., Rogers, S.G., Fraley, R.T. and Beachy, R.N. (1986) Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science* 232, 738-43.

Allen, N.D., Cran, D.G., Barton, S.C., Hettle, S., Reik, W. and Surani, M.A. (1988) Transgenes as probes for active chromosomal domains in mouse development. *Nature* 333, 852-55.

Allen, G., Hall, Jr., G.E., Childs, L.C., Weissinger, A.K., Spiker, S. and Thompson, W.F. (1993) Scafford attachment regions increase reporter gene expression in stably transformed plant cells. *Plant Cell* 5, 603-13.

An, G., Ebert, P.R., Yi, B.-Y. and Choi, C.-H. (1986) Both TATA box and upstream regions are required for the nopaline synthase promoter in transformed tobacco cells. *Mol. Gen. Genet.* 203, 245-50.

Andre, D., Schell, J., Van Montagu, M. and Hernalsteens, J.P. (1986) Gene tagging in plants by a T-DNA insertion that generates APH(3')II plant gene fusions. *Mol. Gen. Genet.* 204, 512-18.

Anzai, H., Yoneyama, K. and Yamaguchi, I. (1989) Transgenic tobacco resistant to a bacterial disease by the detoxification of a pathogenic toxin. *Mol. Gen. Genet.* 219, 492-96.

Baker, B., Schell, J., Lorz, H. and Federoff, N. (1986) Transposition of the maize controlling element *activator* in tobacco. *Proc. Natl. Acad. Sci. USA* 83, 4844-48.

Balcells, L., Swinborne, J. and Coupland, G. (1991) Transposons as tools for the isolation of plant genes. *Trends Biotech.* 9, 31-37.

Baldwin, T.O., Berends, T., Bunch, T.A., Holzman, T.F., Rausch, S.K., Shamansky, L., Treat, M. and Ziegler, M.L. (1984) Cloning of the luciferase

structural genes from *Vibrio harveyi* and expression of bioluminescence in *Escherichia coli*. *Biochemistry* 23, 3663-67.

Bancroft, I., Jones, J.D.G. and Dean, C. (1993) Heterologous transposon tagging of the *DRL1* locus in *Arabidopsis*. *Plant cell* 5, 631-38.

Barnes, W.M. (1990) Variable patterns of expression of luciferase in transgenic tobacco leaves. *Proc. Natl. Acad. Sci. USA* 87, 9183-87.

Barton, K., Binns, A.N., Matzke, A.J.M. and Chilton, M.D. (1983) Regeneration of intact tobacco plants containing full-length copies of genetically engineered T-DNA, and transmission of T-DNA to R1 progeny. *Cell* 32, 1033-43.

Belas, R., Mileham, A., Cohn, D., Hilmen, M., Simon, M. and Silverman, M. (1982) Bacterial bioluminescence: Isolation and expression of the luciferase genes from *Vibrio harveyi*. *Science* 218, 791-92.

Bellen, H.I., O'Kane, C.J., Wilson, C., Grossniklaus, U., Pearson, R.K. and Gehring, W.J. (1989) P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. *Genes & Dev.* 3, 1288-1300.

Benfey, P.N. and Chua, N.-H. (1989) Regulated genes in transgenic plants. *Science* 242, 174-81.

Benfey, P.N., Ren, L. and Chua, N.-H. (1989) The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. *EMBO J.* 8, 2195-2202.

Berk, A.J. and Sharp, P.A. (1977) Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. *Cell* 12, 721-32.

Berry, J.O., Car, J.P. and Klessig, D.F. (1988) mRNAs encoding ribulose-1,5-bisphosphate carboxylase remain bound to polysomes but are not translated

in amaranth seedlines transferred to darkness. *Proc. Natl. Acad. Sci. USA* 85, 4190-94.

Berry, J.O., Breiding, D.E and Klessig, D.F. (1990) Light-mediated control of translational initiation of ribulose- 1,5-bisphosphate carboxylase in amaranth cotyledons. *Plant Cell* 2, 795-803.

Bevan, M., Barnes, W.M. and Chilton, M.D. (1983) Structure and transcription of nopaline synthase gene region of T-DNA. *Nucleic Acids Res.* 11, 369-85.

Bevan, M., Flavell, P.B. and Chilton, M.D. (1983) A chimeric antibiotic resistance gene as a selectable marker for plant cell transformation. *Nature* 304, 184-187.

Bingham, P., Lewis, R. and Rubin, G.M. (1981) Cloning of DNA sequences from the white locus of *D. melanogaster* by a novel and general method. *Cell* 20, 579-95.

Block, A., Dangl, J.L., Hahlbrock, K. and Schulze-Lefert, P. (1990) Functional borders, genetic fine structure, and distance requirements of *cis* elements mediating light responsiveness of the parsley chalcone synthase promoter. *Proc. Natl. Acad. Sci. USA* 87, 5387-91.

Boulton, G.W., Nester, E.W. and Gordon, M.P. (1986) Plant phenolic compounds induce expression of *Agrobacterium tumefaciens* loci needed for virulence. *Science* 232, 983-84.

Bowman, J.L., Smyth, D.R. and Meyerowitz, E.M. (1991) Genetic interactions among floral homeotic genes of *Arabidopsis thaliana*. *Development* 112, 1-20.

Breyne, P., Van Montagu, M., Depicker, A. and Gheysen, G. (1992) Characterization of a plant scafford attachment region in a DNA fragment that normalitzes transgene expression in tobacco. *Plant Cell* 4, 463-71.

Bustos, M.M., Begum, D., Kalkan, F.A., Battraw, M.J. and Hall, T.C. (1991) Positive and negative *cis*-acting DNA domains are required for spatial and temporal regulation of gene expression by a seed storage protein promoter. *EMBO J.* 10, 1469-79.

Callis, J., Fromm, M. and Valbot, V. (1987) Introns increase gene expression in cultured maize cells. *Genes Dev.* 1, 1183-1200.

Caplan, A., Herrera-Estrella, L., Inze, D., Van Montagu, M. and Schell, J. (1983) Introduction of genetic material into plant cells. *Science* 222, 815-21.

Castresana, C., de Carvalho, F., Gheysen, G., Habets, M., Inze, D. and Van Montagu, M. (1990) Tissue-specific and pathogene-induced regulation of a *Nicotiana plumbaginifolia* β-1,3-glucanase gene. *Plant cell* 2, 1131-43.

Castresana, C., Garcia-Luque, I., Alonso, E., Malik, V.S. and Cashmore, A.R. (1988) Both positive and negative regulatory elements mediate expression of photoregulated a CAB gene from *Nicotiana plumbaginifolia*. *EMBO J.* 7, 1929-36.

Christ, C. and Tye, B.K. (1991) Functional domains of the yeast transcription / replication factor MCM1. *Genes & Dev.* 5, 751-63.

Chuck, G., Robbins, T., Nijjar, C., Ralston, E., Courtney-Gutterson, N. and Dooner, H.K. (1993) Tagging and cloning of a petunia flower color gene with the maize transposable element *Activator*. *Plant Cell* 5, 371-78.

Cockerhill, P. and Garrard, W. (1986) Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing toposisomerase II site. *Cell* 44, 941-50.

Coen, E.S. and Meyerowitz, E.M. (1991) The war of the whorls: Genetic interactions controlling flower development. *Nature* 353, 31-37.

Cohn, D.H., Mileham, A.J., Simon, M.I., Nealson, K.H., Rausch, S.K., Bonam, D. and Baldwin, T. (1985) Nucleotide sequence of the *luxA* gene of *Vibrio*

harveyi and the complete amino acid sequence of the alpha subunit of bacterial luciferase. *J. Biol. Chem.* 260, 6139-64.

Comai, L., Dietrich, R.A., Maslyar, D.J., Baden, C.S. and Harada, J. (1989) Coodinate expression of transcriptionally regulated isocitrate lyase and malate synthase genes in *Brassica napus* L. *Plant Cell* 1, 293-300.

Cook, P.R. (1989) The nucleoskeleton and the topology of transcription. *Eur. J. Biochem.* 185, 487-501.

Cooley, L., Kelley, R. and Spradling, A. (1988) Insertional mutagenesis of the *Drosophila* genome with single P elements. *Science* 239, 1121-28.

Coupland, G., Baker, B., Schell, J. and Starlinger, P. (1988) Characterization of the maize transposable element Ac by internal deletions. *EMBO J.* 7, 3653-59.

De Almeida, E.R.P., Gossele, V., Muller, C.G., Dockx, J., Reynaerts, A., Botterman, J., Krebbers, E. and Timko, M.P. (1989) Transgenic expression of two marker genes under the control of an *Arabdopsis rbc* S promoter: Sequences encoding the Rubisco transit peptide increase expression levels. *Mol. Gen. Genet.* 218, 78-86.

De Block, M., Botterman, J., Vandewiele, M., Dockx, J., Thoen, C., Gossele, V., Movva, N.R., Thompson, C., Van Montagu, M. and Leemans, J. (1987) Engineering herbicide resistance in plants by expression of a detoxifying enzyme. *EMBO J.* 6, 2513-18.

De Block, M., Botterman, J., Vandewiele, M., Dockx, J., Thoen, C., Gossele, V., Movva, N.R., Thompson, C., Van Montagu, M. and Leemans, J. (1987) Engineering herbicide resistance in plants by expression of a detoxifying enzyme. *EMBO J.* 6, 2513-18.

De Block, M., Herrera-Estrella, L., Van Montagu, M., Schell, J. and Zambryski, P. (1984) Expression of foreign genes in regenerated plants and their progeny. *EMBO J.* 3, 1681-89.

De Cleene, M. and De Ley, J. (1976) The host range of crown gall. *Bot. Rev.* 42, 389-466.

Dean, C., Favreau, M., Bond-nutter, D., Bedbrook, J. and Dunsmuir, P. (1989) Sequences downstream of translation start regulate quantitive expression of two petunia *rbc*S genes. *Plant Cell* 1, 201-08.

Dean, C., Jones, J., Favreau, M., Dunsmuir, P. and Bedbrook, J. (1988) Influence of flanking sequences on variability in expression levels of an introduced gene in transgenic tobacco plants. *Nucl. Acids Res.* 16, 9267-83.

Deng, X.W., Matsui, M., Wei, N., Wagner, D., Chu, A.M., Feldmann, K.A. and Quail, P.H. (1992) *cop1*, an *Arabidopsis* regulatory gene, encodes a protein with both a Zinc-binding motif and a Gβ homologous domain. *Cell* 71, 791-801.

Diffley, J.F.X. and Stillman, B. (1986) Purification of a cellular double-stranded DNA-binding protein required for initiation of adenovirus DNA replication by using a rapid filter-binding assay. *Mol. Cell Biol.* 6, 1363-73.

Deroles, S.C. and Gardner, R.C. (1988) Expression and inheritance of kanamycin resistance in a large number of transgenic petunias generated by *Agrobactium*-mediated transformation. *Plant Mol. Biol.* 11, 355-64.

Dickey, L.F., Gallo-Meagher, M. and Thompson, W.F. (1992) Light regulatory sequences are located within the 5' protion of the *Fed-1* massage sequence. *EMBO J.* 11, 2311-17.

Dillon, N. and Grosveld, F. (1993) Transcriptional regulation of multigene loci: multilevel control. *TIG.* 9, 134-37.

Donald, R.G.K. and Cashmore, A.R. (1990) Mutation of either G box or I box sequences profoundly affects expression from the *Arabidopsis rbc*S-1A promoter. *EMBO J.* 9, 1717-26.

Douglas, C.J., Staneloni, R.J., Rubin, R.A. and Nester, E.W. (1985) Identification and genetic analysis of an *Agrobacterium tumefaciens* chromosomal virulence region. *J. Bact.* 161, 850-60.

Draper, J., Scott, R. and Armitage, P. (1989) Plant genetic transformation and gene expression. A laboratory manual. Blackwell Scientific Publications, Oxford. 199-208.

Drews, G.N. and Goldberg, R.B. (1989) Genetic control of flower development. *TIG.* 5, 256-61.

Dubois, E., Bercy, J., Descamps, F. and Messenguy, F. (1987) Characterization of two new genes essential for vegetative growth in *Saccharomyces cerevisiae*, nucleotide sequence determination and chromosome mapping. *Gene* 55, 265-75.

Dynan, W.S. (1989) Modularity in promoters and enhancers. *Cell* 58, 1-4

Elliston, K. and Messing, J. (1988) The molecular architecture of plant genes: A phylogenetic perspective. In: Architecture of Eukaryotic genes, Kahl, G (ed), pp 21-56, VCH Publisher, New York.

Escher, A.P., O'Kane, D.J., Lee, J. and Szalay, A.A. (1989) Bacterial luciferase $\alpha\beta$ fusion protein is fully active as a monomer and highly sensitive *in vivo* to elevated temperature. *Proc. Natl. Acad. Sci. USA* 86, 6528-32.

Ettinger, W.F. and Harada, J.J. (1990) Translational or post-translational processes affect differentially the accumulation of isocitrate lyase and malate synthesis proteins and enzyme activities in embryo and seedlings of *Brassica napus*. *Arch. Biochem. Biophys.* 281, 139-43.

Fang, R.-X., Nagy, F., Sivasubramaniam, S. and Chua, N.-H. (1989) Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell* 1, 141-50.

Fedoroff, N., Wessler, S. and Shure, M. (1983) Isolation of the transposable maize controlling elements Ac and Ds. *Cell* 35, 235-42.

Fedoroff, N.V., Furtek, D.B. and Nelson, O.E. (1984) Cloning of the *bronze* locus in maize by a simple and generalizable procedure using the transposable element Activator (Ac). *Proc. Natl. Acad. Sci. USA* 81, 3825-29.

Feldmann, K.A. (1991) T-DNA insertion in *Arabidopsis*: mutational spectrum. *Plant J.* 1, 71-82.

Feldmann, K.A., Marks, M.D. (1986) Rapid and efficient regeneration of plants from explants of *Arabidopsis thaliana*. *Plant Sci.* 47, 63-69.

Feldmann, K.A., Marks, M.D. (1987) *Agrobacterium*-mediated transformation of germinating seeds of *Arabidopsis thaliana*: anon-tissue culture approach. *Mol. Gen. Genet.* 208, 1-9.

Feldmann, K.A., Marks, M.D., Christianson, M.L. and Quatrano, R.S. (1989) A dwarf mutant of *Arabidopsis* generated by T-DNA insertion mutagenesis. *Science* 243, 1351-54.

Forget, B.G. (1993) YAC transgenes: Bigger is probably better. *Proc. Natl. Acad. Sci. USA* 90, 7909-11.

Fraley, R.T., Rogers, S.G., Horsch, R.B., sanders, P.R., Flick, J.S., Adams, S.P., Bittner, M.L., Brand, L.A., Fink, C.L., Fry, J.S., Galluppi, G.R., Goldberg, S.B., Hoffmann, N.L and Woo, S.C. (1983) Expression of bacterial genes in plant cells. *Proc. Natl. Acad. Sci. USA* 80, 4803-7.

Fried, M. and Crothers, D.M. (1981) Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* 9, 6505-29.

Futterer, J., Gordon, K., Sanfacon, H., Bonneville, J.-M. and Hohn, T. (1990) Positive and negative control of translation by the leader sequence of cauliflower mosaic virus pregenomic 35S RNA. *EMBO J.* 9, 1697-1707.

Galas, D.J. and Schmitz, A. (1978) DNAase footprinting: a simple mehtod for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* 5, 3157-70.

Gallie, D.R., Feder, J.N., Schimke, R.T. and Walbot, V. (1991) Post-transcriptional regulation in higher eukaryotes: The role of the reporter gene in controlling expression. *Mol. Gen. Genet.* 228, 258-64.

Garner, M.M. and Revzin, A. (1981) A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of Escherichia coli operon regulatory system. *Nucleic Acids Res.* 9, 3047-3060.

Gasser, C.S. and Fraley, R.T. (1989) Genetically engineering plants for crop improvement. *Science* 244, 1293-99.

Gasser, S.M. and Laemmli, U.K. (1986) Cohibitation of scaffold binding regions with upstream enhancer elements of three developmentally regulated genes of *D. melanogaster*. Cell 46, 521-30.

Gheysen, G., Villarroel, R. and Van Montagu, M. (1991) Illegitimate recombination in plants: a model for T-DNA integration. *Genes & Dev.* 5, 287-97.

Gidoni, D., Bond-Nutter, D., Brosio, P., Jones, J., Bedbrook, J. and Dunsmuir, P. (1988) Coordinated expression between two photosynthetic petunia genes in transgenic plants. *Mol. Gen. Genet.* 211, 507-14.

Gidoni, D., Brosio, P., Bond-Nutter, D., Bedbrook, J. and Dunsmuir, P. (1989) Novel *cis*-acting elements in petunia *cab* gene promoters. *Mol. Gen. Genet.* 215, 337-44.

Gierl, A. and Saedler, H. (1992) Plant -transposable elements and gene tagging. *Plant Mol. Biol.* 19, 39-49.

Gilmartin, P.M., Sarokin, L., Memelink, J. and Chua, N.-H. (1990) Molecular light switches for plant genes. *Plant cell* 2, 369-78.

Goodman, M.A. (1988) The chromatin domain as unit of gene regulation. *BioEssays* 9, 50-55.

Goldsbrough, A. and Bevan, M. (1991) New patterns of gene activity in plants detected using an *Agrobacterium* vector. *Plant Mol. Biol.* 16, 263-69.

Greenwald, I. (1985) *Lin*-12, a nematode homeotic gene, is homologous to a set of mammalian proteins that includes epidermal growth factor. *Cell* 43, 583-90.

Gridley, T., Soriano, P. and Jaenisch, R. (1987) Insertional mutagenesis in mice. *Trends Genet.* 3, 162-66.

Grosveld, F., Blom van Assendelft, G., greaves, D. and Kollias, G. (1987) Position-independent, high-level expression of the human ß Globin gene in transgenic mice. *Cell* 51, 975-85.

Guarente, L. (1988) UASs and enhancers: Common mechanism of transcriptional activation in yeast and mammals. *Cell* 52, 303-05.

Hastings, W.J. and Nealson, K.H. (1977) Bacterial bioluminescence. *Annu. Rev. Microbiol.* 31, 549-95.

Hayashi, H., Czaja, I., Lubenow, H., Schell, J. and Walden, R. (1992) Activation of a plant gene by T-DNA tagging: auxin-independent growth *in vitro*. *Science* 258, 1350-53.

Herman, L., Jacobs, A., Van Montagu, M. and Depicker, A. (1990) Plant chromosome/marker gene fusion assay for study of normal and truncated T-DNA integration events. *Mol. Gen. Genet.* 224, 248-56.

Herrera-Estrella, A., Chen, Z., Van Montagu, M. and Wang, K. (1988) Vir D proteins of *Agrobactium tumefaciens* are required for the formation of a

covalent DNA-protein complex at the 5' terminus of T-strand molecules. *EMBO J.* 7, 4055-62.

Herrera-Estrella, L., De Block, M., Messens, E., Hernasteens, J.P., Van Montagu, M. and Schell, J. (1983a) Chimeric genes as dominant selectable markers in plant cells. *EMBO J.* 2,987-95.

Herrera-Estrella, L., Depicker, A., Van Montagu, M. and Schell, J. (1983b) Expression of chimeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature* 303, 209-13.

Hightower, R., Baden, C., Penzes, E., Lund, P. and Dunsmuir, P. (1991) Expression of antifreeze proteins in transgenic plants. *Plant Mol. Biol.* 17, 1013-21.

Hille, J., Van Kan, J., Klasen, I. and Schilperoort, R. (1983) Site-directed mutagenesis in *Escherichia coli* of a stable R772::Ti cointegrate plasmid from *Agrobacterium tumefaciens*. *J. Bacteriol.* 154, 693-701.

Hobbs, S.L.A., Kpodar, P. and DeLong, C.M.O. (1990) The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. *Plant Mol. Biol.* 15, 851-64.

Hobbs, S.L.A., Warkentin, T. and DeLong, C.M.O. (1993) Transgene copy number can be positively or negatively associated with transgene expression. *Plant Mol. Biol.* 21, 17-26.

Hoekma, A., Hirsch, P.R., Hooykaas, P.J. and Schilperoort, R.A. (1983) A binary plant vector strategy based on the separation of the *vir* and T-region of *Agrobacteria*. *Nature* 303, 1051-55.

Hoekma, A., Roelvink, P., Hooykaas, P.J. and Schilperoort, R.A. (1984) Delivery of T-DNA from the *Agrobacterium tumefaciens* chromosome into plant cells. *EMBO J.* 3, 2485-90.

Hohn, B., Koukolikova-Nicola, Z., Durrenberger, F., Bakkeren, G. and Koncz, C. (1991) The T-DNA on its way from *Agrobacterium tumefaciens* to the plants. *In Advances in Molecular Genetics of Plant-Microbe Interactions.* 1, 19-27.

Hooykaas, P.J.J., Klapwijk, P.M., Nuti, M.P., Schilperoort, R.A. and Rorsch, A. (1977) Transfer of the *Agrobacterium tumefaciens* Ti plasmid to avirulent agrobacteria and to *Rhizobium explanta. J. Gen. Microbiol.* 98, 477-84.

Hooykaas, P.J.J. and Schilperoort, R.A. (1992) *Agrobacterium* and plant genetic engineering. *Plant Mol. Biol.* 19, 15-38.

Hope, I.A. (1991) Promoter trapping in *Caenorhabditis elegans. Development* 113, 571-4.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. (1985) A simple and general method for transferring genes into plants. *Science* 227, 1229-31.

Howard, A.E., Winsor, B.A., De Vos, G. and Zambryski, P. (1989) Activation of the T-DNA transfer process in *Agrobacterium* results in the generation of a T-strand-protein complex: tight association of VirD2 with the 5' ends of T-strands. *Proc. Natl. Acad. Sci. USA* 86, 4017-21.

Howard, A.E., Zupan, J.R. and Zambryski, P. (1992) The Vir D2 protein of *Agrobacterium tumefaciens* contains a C-terminal bipartite localization signal: implications for nuclear uptake of DNA in plant cells. *Cell* 68, 109-18.

Ingelbrecht, I.L., Herman, L.M.F., Dekeyser, R.A., Van Montagu, M. and Depicker, A.G. (1989) Different 3' end regions strongly influence the level of gene expression in plant cells. *Plant Cell* 1, 381-90.

Itakura, K., Rossi, J.J. and Wallace, R.B. (1984) Synthesis ane use of synthetic oligonucleotides. *Annu. Rev. Biochem.* 53, 323-56.

Jack, T., Brockman, L.L. and Meyerowitz, E.M. (1992) The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* 68, 683-97.

Jefferson, R.A. (1987) Assaying chimeric genes in plants: GUS gene fusion system. *Plant Mol. Biol. Rep.* 5, 387-405.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M. (1987) GUS fusion: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6, 3901-7.

Jiang, C, Escher, A. and Szalay, A.A. (1993a). Alteration of a luciferase based binary plant transformation vector for simultaneous expression of desired gene products and bacterial luciferase in transgenic plants. Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds.) John Wiley & Sons. 127-31.

Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Identification of plant genes *in vivo* by tagging with T-DNA border-linked luciferase genes followed by inverse polymerase chain reaction amplification. *Plant Mol. Biol. Rep.* 10, 345-61.

Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1993b) Isolation of developmentally regulated promoter elements from tobacco by T-DNA insertional activation with promoterless luciferase marker gene. Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds) John Wiley & Sons. 217-21

Jones, J.D.G., Gilbert, D.E., Grady, K.L. and Jorgensen, R.A. (1987) T-DNA structure and gene expression in petunia plants transformed by *Agrobacterium tumefaciens* C58 derivatives. *Mol. Gen. Genet.* 207, 478-85.

Jordan, E., Saedler, H. and Starlinger, P. (1968) 0^o and strong polar mutations in the gel operon are insertions. *Mol. Gen. Genet.* 102, 701-07.

Jorgensen, R. (1990) Altered gene expression in plants duo to trans interactions between homologous genes. *Trends Biotechnol.* 8, 340-44.

Jorgensen, R. (1991) Beyond antisense-How do transgenes interact with homologous plant genes? *Trends Biotechnol.* 9, 266-67.

Jorgensen, R., Snyder, C. and Jones, J.D.G. (1987) T-DNA is organized predominantly in reverted repeat structures in plants transformed with *Agrobacterium tumefaciens* C58 derivatives. *Mol. Gen. Genet.* 207, 471-477.

Kado, C.I. (1991) Molecular mechanisms of crown gall tumorigenesis. *Crit. Rev. Plant Sci.* 10, 471-77.

Kadonaga, J.T. and Tjian, R. (1986) Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. USA* 83, 5889-93.

Kahl, G. and Schell, J. (1982) Molecular Biology of Plant Tumours. New York, Academic Press.

Kanemoto, R.H., Powell, A.T., Akiyoshi, D.E., Regier, D.A., Kerstetter, R.A., Nester, E.W., Hawes, M.C. and Gordon, M.P. (1989) Nucleotide sequence and analysis of the plant-inducible locus *pinF* from *Agrobacterium tumefaciens*. *J. Bacteriol.* 171, 2506-12.

Keller, B., Schmid, J. and Lamb, C. J. (1989) Vascular expression of a bean cell wall glycin-rich protein-β-glucuronidaes gene fusion in transgenic tobacco. *EMBO J.* 8, 1309-14.

Kertbundit, S., De Greve, H., DeBoeck, F., Van Montagu, M. and Hernalsteens, J.P. (1991) *In vivo* random β-glucuronidase gene fusion in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 88, 5212-16.

Kieber, J.J., Rothenberg, M., Roman, G., Feldmann, K.A. and Ecker, J.R. (1993) *CTR1*, a negative regulator of ethylene response pathway in *Arabidopsis*, encodes a member of the *Raf* family of protein kinases. *Cell* 72, 427-441.

Kleckner, N. (1981) Transposable elements in prokaryotes. *Annu. Rev. Genet.* 15, 341-404.

Kleckner, N., Roth, J. and Botstein, D. (1977) Genetic engineering *in vivo* using translocatable drug-resistance elements. *J. Mol. Biol.* 116, 125-129.

Koncz, C., Nemeth, K., Redei, G. and Schell, J. (1992) T-DNA insertional mutagenesis in *Arabidopsis*. *Plant Mol. Biol.* 20, 963-976.

Koncz, C., Langridge, W.H.R, Olsson, O, Schell, J. and Szalay, A.A. (1990) Bacterial and firefly luciferase Genes in transgenic plants: advantages and disadvantages of a reporter gene. *Devel. Genetics* 11, 224-32.

Koncz, C., Olsson, O., Langridge, W.H.R., Schell, J. and Szalay, A.A. (1987) Expression and assembly of functional bacterial luciferase in plants. *Proc. Natl. Acad. Sci. USA* 84, 131-35.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Korber, H., Redei, G.P. and Schell, J. (1989) High-frequency T-DNA-mediated gene tagging in plants. *Proc. Natl. Acad. Sci. USA* 86, 8467-71.

Koncz, C., Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Reiss, B., Redei, G.P. and Schell, J. (1990) Isolation of a gene encoding a novel chloroplast protein by T-DNA tagging in *Arabidopsis thaliana*. *EMBO J.* 9, 1337-46.

Koncz, C. and Schell, J. (1986) The promoter of the T1-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* 204, 383-96.

Koncz, C., Schell, J. and Redei, G.P. (1992) T-DNA transformation and insertion mutagenesis. In: Methods in *Arabidopsis* Research. Koncz, C., Chua, N.H. and Schell, J. (Eds.) World Scientific. 224-73.

Kuhlemeier, C. (1992) transcriptional and post-transcriptional regulation of gene expression in plants. *Plant Mol. Biol.* 19, 1-14.

Kuhlemeier, C., Fluhr, R., Green, P.J. and Chua, N.-H. (1987) Sequences in the pea *rbc*S-3A gene have homology to constitutive mammalian enhancers but function as negative regulatory elements. *genes Dev.* 1, 247-55.

Kuhlemeier, C., Strittmatter, G., Ward, K. and Chua, N.-H. (1989) The pea *rbc*S-3A promoter mediates light responsiveness but not organ specificity. *Plant Cell* 1, 471-78.

Kurfurst, M., Ghisla, S. and Hastings, W. (1984) Characterization and postulated structure of the primary emitter in the bacterial luciferase reaction. *Proc. Natl. Acad. Sci. USA* 81, 2990-94.

Lam, E., Benfey, P.N., Gilmartin, P.M., Fang, R.-X. and Chua, N.-H. (1989) Site-specific mutations alter *in vitro* factor binding and change promoter expression pattern in transgenic plants. *Proc. Natl. Acad. Sci. USA* 86, 7890-94.

Lam, E and Chua, N.-H. (1990) GT-1 binding site confers light responsive expression in transgenic tobacco. *Science* 248, 471-74.

Lamb, C.J., Ryals, J.A., Ward, E.R. and Dixon, R.A. (1992) Emerging strategies for enhancing crop resistance to microbial pathogens. *Biotechnology* 10, 1436-46.

Langridge, W.H.R., Escher, A.P. and Szalay, A.A. (1991) Measurement of bacterial luciferase as a reporter enzyme *in vivo* in transformed bacteria, yeast and plant cells and in transgenic plants. *Technique* 3, 99-108.

Langridge, W.H.R., Fitzgerald, K.J., Koncz, C., Schell, J. and Szalay, A.A. (1989) Dual promoter of *Agrobacterium tumefaciens* mannopine synthase genes is regulated by plant growth hormones. *Proc. Natl. Acad. Sci. USA* 86, 3219-23.

Leutwiler, L.S., Hough-Evans, B.R. and Mayerowitz, E.M. (1984) The DNA of *Arabidopsis thaliana*. *Mol. Gen. Genet.* 194, 15-23.

Levis, R., O'Hare, K. and Rubin, G.M. (1984) Effects of transposable element insertions on RNA encoded by the *white* gene of *Drosophila*. *Cell* 38, 471-481.

Lindsey, K., Wei, W., Clarke, M.C., McArdle, H.F., Rooke, L.M. and Topping, J.F. (1993) Tagging genomic sequences that direct transgene expression by activation of a promoter trap in plants. *Transgenic Res.* 2, 33-47.

Mariani, C., De Beuckeleer, M., Truettner, J., Leemans, J. and Goldberg, R.B. (1990) Induction of male sterility in plants by a chimeric ribonuclease gene. *Nature* 347, 737-41.

Maniatis, T., Goodbourn, S. and Fischer, J.A. (1987) Regulation of inducible and tissue-specific gene expression. *Science* 236, 1237-45.

Marcotte, W.R., Russell, S.H and Quatrano, R.S. (1989) Abscisic acid-responsive sequences from the Em gene of wheat. *Plant Cell* 1, 969-76.

Marks, M.D. and Feldmann, K,A. (1989) Trichome development in *Arabidopsis thaliana*. I. T-DNA tagging of the *GLABROUS1* gene. *Plant Cell* 1, 1043-1050.

Mason, H.S., Lam, D.M.K. and Arntzen, C. (1992) Expression of hepatitis B surface antigen in transgenic plants. *Proc. Natl. Acad. Sci. USA* 89, 11745-49.

Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Bakkeren, G., Crameri, A., Angelis, K., Redei, G.P., Schell, J., Hohn, B. and Koncz, C. (1991) T-DNA integration: a model of illegitimate recombination in plants. *EMBO J.* 10, 697-704.

Mayerowitz, E.M. (1987) *Arabidopsis thaliana*. *Annu. Rev. Genet.* 21, 93-111.

McClintock, B. (1951) Chromosome organization and genetic expression. *Cold Spring Harbor Symp. Quant. Biol.* 16, 13-47.

McKendree, W.L., Paul, A.-L., DeLisle, A.J. and Ferl, R.J. (1990) *In vivo* and *in vitro* characterization of protein interactions with the dyed G-box of the Arabidopsis *Adh* gene. *Plant Cell* 2, 207-14.

McKnight, S.L., Gavis, E.R. and kingsbury, R. (1981) Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upsteam control region. *Cell* 25, 385-98.

Meyerowitz, E.M. (1992) Introduction to the *Arabidopsis* genome. In: Methods in *Arabidopsis* Research. Koncz, C., Chua, N.H. and Schell, J. (Eds.) World Scientific. 100-118.

Meyerowitz, E.M., Smith, D.R. and Bowman, J.L. (1989) Abnormal flowers and pattern formation in floral development. *Development* 106, 209-17.

Meyerowitz, E.M., Bowman, J.L., Brockman, L.L., Dews, G.N., Jack, T., Sieburth, L.E. and Weigel, D. (1991) A genetic and molecular model for flower development in *Arabidopsis thaliana*. *Development Supplement* 113, 153-63.

Michel, D., Salamini, F., Bartels, D., Dale, P., Baga, M. and Szalay, A.A. (1993) Analysis of a desiccation and ABA-responsive promoter isolated from the resurrection plant *Craterostigma plantagineum*. *Plant J.* 4, 29-40.

Murai, N., Li, Z., Kawagoe, Y. and Hayashimoto, A. (1991) Transposition of the maize activator element in transgenic rice plants. *Nucleic Acid Res.* 19, 617-22.

Newman, C.T., Ohme-Takagi, M., Taylor, C.B. and Green, P.L. (1993) DST sequences, highly conserved among plant SAUR genes, target reporter transcripts for rapid decay in tobacco. *Plant Cell* 5, 701-04.

Norman, C., Runswick, M., Pollock, R. and Treisman, R. (1988) Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-*fos* serum response element. *Cell* 55, 989-1003.

Ochman, H., Gerber, A.S. and Hartl, L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* 120, 621-23.

Odell, J.T., Nagy, F. and Chua, N.-H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313, 810-12.

Okamuro, J.E. and Goldberg, R.B. (1985) Tobacco single-copy DNA is highly homologous to sequences present in genomes of its diploid progenitors. *Mol. Gen. Genet.* 108, 290-98.

Okamuro, J.E. and Goldberg, R.B. (1989) Regulation of plant gene expression: General principles. In: The Biochemistry of Plants, Vol. 15, A. Marcus (ed), Academic Press, San Diego, 1-81.

Olsson, O., Escher, A.P., Sandberg, G., Schell, J., Koncz, C. and Szalay, A.A. (1989) Engineering of monomeric bacterial luciferases by fusion of *luxA* and *luxB* genes in *Vibrio harveyi*. *Gene* 81, 335-47.

Ondek, B., Gloss, L. and Herr, W. (1988) The SV40 enhancer contains two distinct levels of organization. *Nature* 333, 40-45.

Peach, C. and Velten, J. (1991) Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. *Plant Mol. Biol.* 17, 19-60.

Perlak, F.J., Fuchs, R.L., Dean, D.A., McPherson, S.L. and Fischhoff, D.A. (1991) Modification of the coding sequence enhances plant expression of insect control protein genes. *Proc. Natl. Acad. Sci. USA* 88, 3324-28.

Ream, W. (1990) *Agrobacterium tumefaciens* and interkingdom genetic exchange. *Annu. Rev. Phytopathol.* 27, 583-618.

Reiss, B., Sprengel, R., Will, H. and Schaller, H. (1984) A new sensitive method for qualitative and quantitative analysis of neomycin phosphotransferase in crude cell extracts. *Gene* 30, 211-18.

Roeder, G.S., Farabaugh, P.J., Chaleff, D.F. and Fink, G.R. (1980) The origins of gene instability in yeast. *Science* 209, 1375-80.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350-54.

Scheid, O.M., Paszkowski, J. and Potrykus, I. (1991) Reversible inactivation of a transgene in *Arabidopsis thaliana*. *Mol. Gen. Genet.* 228, 104-112.

Schell, J. (1987) Transgenic plants as tools to study the molecular organization of plant genes. *Science* 237,1176-82.

Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. and Sommer, H. (1990) Genetic control of flower development by homeotic genes in *Antirrhinum majus*. *Science* 250, 931-36.

Schwarz-Sommer, Z., Hue, I., Huijser, P., Flor, P.J., Hansen, R., Tetens, F., Lonnig, W.E., Saedler, H. and Sommer, H. (1992) Characterization of the *Antirrhinum* floral homeotic MADS-box gene *DEFICIENS*: Evidence for DNA binding and autoregulation of its persistent expression throughout flower development. *EMBO J.* 11, 251-63.

Sheen, J. (1991) Molecular mechanisms underlying the differential expression of maize pyruvate, orthophosphate dikinase genes. *Plant Cell* 3, 225-45.

Shortle, D. and Botstein, D. (1983) Directed mutagenesis with sodium bisulfite. *Methods Enzymol.* 100, 457-509.

Singh, J.H., LeBowitz, J.H., Baldwin Jr., A.S. and Sharp, P.A. (1988) Molecular cloning of an enhancer binding protein: isolation by screening of expression library with a recognition site DNA. *Cell* 52, 415-23.

Smith, C.J.S., Watson, C.F., Ray, J., Bird, C.R., Morris, P.C., Schuch, W. and Grierson, D. (1988) Antisense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes. *Nature* 334, 724-26.

Smith, E.F. and Townsend, C.O. (1907) A plant tumor of bacterial origin. *Science* 25, 671-73.

Smith, M. (1985) *In vitro* mutageneiss. *Annu. Rev. Genet.* 19, 423-63.

Sommer, H., Beltran, J.P., Huijser, P., Pape, H., Lonnig, W.E., saedler, H. and Schwarz-Sommer, Z. (1990) *deficiens*, a homeotic gene involved in control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. *EMBO J.* 9, 877-92.

Sommer, H., Carpenter, R., Harrison, B.J. and Saedler, H. (1985) The transposable element Tam3 of *Antirrhinum majus*: generates a novel type of sequence alterations upon excision. *Mol. Gen. Genet.* 199, 225-31.

Southern, E. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98, 503-18.

Stachel, S.E., Messens, E., Van Montagu, M. and Zambryski, P. (1985) Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. *Nature* 318, 624-69.

Stachel, S.E. and Nester, E.W. (1986) The genetic end transcriptional organization of the *vir* region of the A6 Ti plasmid of *Agrobacterium tumefaciens*. *EMBO J.* 7, 1445-54.

Stachel, S.E. .and Zambryski, P. (1986) *Vir A* .and *vir G* control the plant-induced activation of the T-DNA transfer process of *Agrobacterium tumefaciens*. *Cell* 46, 325-33.

Starlinger, P. and Saedler, H. (1972) Insertion mutations in microorganisms. *Biochemie* 54, 177-85.

Stougaard, J., Sandal, N.N., Grøn, Khule, A. and Marcker, K.A. (1987) 5' analysis of the soybean leghaemoglobin *lbc*3 gene: Regulatory elements required for promoter activity and organ specificity. *EMBO J.* 6, 3565-69.

Struhl, K. (1987) Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell* 49, 295-97.

Sullivan, M. and Green, P.J. (1993) Post-transcriptional regulation of nuclear-encoded genes in higher plants: the roles of mRNA stability and translation. *Plant Mol. Bol.* 23, 1091-1104.

Taylor, J.L., Jones, J.D.G., Sandler, S., Mueller, G.M., Bedbrook, J. and Dunsmuir, P. (1987) Optimizing the expression of chimeric genes in plant cells. *Mol. Gen. genet.* 210, 572-77.

Teeri, TH., Herrera-Estrella, L., Depicker, A., van Montagu, M and Palva, ET. (1986) Identification of plant promoters in situ by T-DNA-mediated transcriptional fusions to the *npt II* gene. *EMBL J.* 5, 1755-60.

Thornburg, R.W., An, G., Cleveland, T.E., Johnson, R. and Ryan, C.A. (1987) Wound-inducible expression of a potato inhibitor II-chloraphenicol acetyltransferase gene fusion in transginc tobacco plants. *Proc. Natl. Acad. Sci. USA* 84, 744-48.

Topping, J.F., Wei, W. and Lindsey, K. (1991) Functional tagging of regulatory elements in the plant genome. *Development* 112, 1009-19.

Trobner, W., Ramirez, L., Motte, P., Hue, I., Huijser, P., Lonnig, W.E., Saedler, H., Sommer, H. and Schwarz-Sommer, Z. (1992) *GLOBOSA*: A homeotic gene which interacts with *DEFICIENS* in the control of *Antirrhinum* floral organogenesis. *EMBO J.* 11, 4693-4704.

van der Krol, A.R., Lenting, P.E., Veenstra, J., van der Meer, I.M., Koes, R.E., Gerats, A.G.M., Mol, J.N.M. and Stuitje, A.R. (1988) An anti-sense chalcone

synthase gene in transgenic plant inhibits flower pigmentation. *Nature* 333, 866-69.

Van Dyke, K. (1985) Bioluminescence and Chemiluminescence: Instruments and Applications. Boca Raton, Florida, CRC Press.

Van Haute, E., Joos, H., Maes, M., Warren, G., Van Montagu, M. and Schell, J. (1983) Intergeneric transfer and exchange recombination of restriction fragments cloned in pBR322: a novel strategy for reversed genetics of the Ti plasmid of *Agrobacterium tumefaciens*. *EMBO J.* 2, 411-17.

van Larebeke, N., Genetello, C., Schell, J., Schilperoort, R.A., Hermans, A.K., Hernalsteens, J.P. and Van Montagu, M. (1975) Acquisition of tumor-inducing ability by non-oncogenic agrobacteria as a result of plasmid transfer. *Nature* 255, 742-43.

Van Sluys, M.A., Tempe, J. and Fedoroff, N. (1987) Studies on the introduction and mobility of maize *Activator* element in *Arabidopsis thaliana* and *Daucus carota*. *EMBO J.* 6, 3881-89.

Van Veen, R.J.M., den Dulk-Ras, H., Bisseling, T., Schilperoort, R.A. and Hooykaas, P.J.J. (1988) Grown gall tumor and root nodule formation by the bacterium *Phyllobacterium myrsinacearum* after the introduction of an *Agrobacterium* Ti plasmid or a *Rhizobium* Sym plasmid. *Mol. Plant-Microbe Interact.* 1, 231-34.

Visser, R.G.F., Somhorst, I., Kuipers, G.J., Ruys, N.J., Feenstra, W.J. and Jacobsen, E. (1991) Inhibition of the expression of the gene for granule-bound starch synthase in potato by antisense constructs. *Mol. Gen. Genet.* 225, 289-96.

Von Schaewen, A., Stitt, M., Schmidt, R., Sonnewald, U. and Willmitzer, L. (1990) Expression of a yeast-derived invertase in the cell wall of tobacco and *Arabidopsis* plants leads to accumulation of carbohydrate and inhibition of

photosynthesis and strongly influences growth and phenotype of transgenic tobacco plants. *EMBO J.* 9, 3033-44.

Walbot, V. (1992) Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis. *Annu. Rev. Plant Phhysiol. Plant Mol. Biol.* 43, 49-82.

Walden, R., Koncz, C. and Schell, J. (1990) The use of gene vectors in plant molecular biology. *Methods Mol. Cell Biol.* 1, 175-94.

Walden, R., Hayashi, H. and Schell. J. (1991) T-DNA as a gene tag. *Plant J.* 1, 281-288.

Waldron, C., Murphy, E.B., Roberts, J.L., Gustafson, G.D., Armour, S.L. and Malcolm, S.K. (1985) Resistance to hygromycin B. A marker for plant transformation. *Plant. Mol. Biol.* 5, 103-8.

Wasylyk, B. (1988) Enhancers and transcription factors in control of gene expression. *Biochimica et Biophysica Acta* 951, 17-35.

Weising, K., Schell, J. and Kahl, G. (1988) Foreign genes in plants: transfer structure and expression. *Annu. Rev. Genet.* 22, 421-77.

Williamson, V.M., Young, E.T. and Ciriacy, M. (1981) Transposable elements associated with constitutive expression of yeast alcohol dehydrogenase II. *Cell* 23, 11-17.

Wilson, C., Pearson, R.K., Bellen, H.J., O'Kane, CJ., Grossniklaus, U. and Gehring, W.J. (1989) P-element-mediated enhancer detection: an efficient method for isolation and characterizing developmentally regulated genes in *Drosophila. Genes & Dev.* 3, 1201-1213.

Wu, C. (1985) An exonuclease protection assay reveals heat-shock element and TATA box DNA-binding proteins in crude nuclear extracts. *Nature* 317, 84-87.

Yadav, N.S., Vanderleyden, J., Bennett, D.R., Barnes, W.M. and Chilton, M.D. (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc. Natl. Acad. Sci. USA* 79, 6322-26.

Yamaguchi-Shinozaki, K. and Shinizaki, K. (1994) A novel *cis*-acting element in an *Arabidopsis* gene is invilved in responsiveness to drought, low-tempererature, or high-salt stress. *Plant Cell* 6, 251-64.

Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A. and Meyerowitz, E.M. (1990) The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature* 346, 35-39.

Yoder, J.I., Palys, J., Alpert, K. and Lassner, M. (1988) *Ac* transposition in transgenic tomato plants. *Mol. Gen. Genet.* 213, 291-96.

Youn, C. and Nester, E. W. (1988) Association of the Vir D protein with 5' end of T-strand in *Agrobacterium tumefaciens*. *J. Bacteriol.* 170, 3367-74.

Zaenen, I., Van Larebeke, N., Teuchy, H., Van Montagu, M. and Schell, J. (1974) Supercoiled circular DNA in crown-gall inducing *Agrobacterium* strains. *J. Mol. Biol.* 86, 109-27.

Zambryski, P. (1988) Basic processes underlying *Agrobacterium*-mediated DNA transfer to plant cells. *Annu. Rev. Genet.* 22, 1-30.

Zambryski, P. (1992) Chronicles from the *Agrobacterium*-plant cell DNA transfer story. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 43, 465-90.

Zambryski, P., Joos, H., Genetello, C., Leemans, J., Van Montagu, M. and Schell, J. (1983) Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. *EMBO J.* 2, 2143-50.

Zambryski, P., Tempe, J. and Schell, J. (1989) Transfer and function of T-DNA genes from *Agrobacterium* Ti and Ri plasmids in plants. *Cell* 56, 193-201.

Ziegler, M.M. and Baldwin, T.O. (1981) Biochemistry of bacterial bioluminescence. Curr. Top. Bioenerg. 12, 65-113.

Chapter III

Alteration of a Luciferase Based Binary Plant Transformation Vector for Simultaneous Expression of Desired Gene Products and Bacterial Luciferase in Transgenic Plants

INTRODUCTION

Agrobacterium tumefaciens-mediated gene transfer is a commonly used method to introduce genes into cells of higher plants and to study gene regulation in regenerated transformants (Gasser and Fraley, 1989; Schell, 1987). The plant expression vector pPCV701 was constructed (Koncz *et al.*, 1987) to transfer a gene of choice from *A. tumefaciens* into plant cells. This vector contains a *ColE1* origin of DNA replication, ampicillin resistance and *npt II* genes (for kanamycin resistance) as positive selection markers for transformation experiments. In addition, this vector contains bidirectional mannopine synthase (mas) promoters (Velten *et al.*, 1984; Koncz *et al.*, 1987). Previous studies have shown that the bidirectional mas (P1, P2) promoters can be used to control the expression of two genes simultaneously in higher plants and that their expression patterns are regulated by plant hormones (Langridge *et al.*, 1989). To monitor gene expression from the mas promoters in plant cells *in vivo*, a *luxA&B* luciferase gene fusion (*fab2*) was used as a marker gene (Escher *et al.*, 1989). The *fab2* gene encodes a monomeric luciferase enzyme that catalyses emission of visible light *in vivo* upon addition of the substrate decanal (Koncz *et al.*, 1990). The *fab2* gene was transcriptionally fused to the mas P1 or P2 promoters by cloning the *fab2* DNA fragment into the Sal I site or the Bam HI site, respectively. Luciferase activity was detected *in vivo* using the Hamamatsu Argus 100 low-light video image analyzer (Langridge *et al.*, 1991; Jiang *et al.*, 1992). Bioluminescence, therefore, offers a non-destructive and sensitive means for monitoring regulation of gene expression in higher plants (Koncz *et al.*, 1990).

There are, however, several drawbacks when this vector is used for cloning. First, there is only one restriction site available for each promoter (Bam HI for P2; Sal I for P1) (Koncz *et al.*, 1987). Second, the DNA fragment of

the *fab2* reporter gene contains a restriction site identical to the cloning sites in the vector. Therefore, cutting at the cloning sites in the vector will cleave the reporter gene as well.

In order to overcome these drawbacks and make this vector more convenient for the study of gene regulation and expression in higher plants, the DNA fragment of the *fab2* gene was modified by removing the sites identical to the cloning sites present in the vector pPCV701. The new vector can now be used to clone a DNA fragment without cleavage of the reporter gene. In addition, a multiple cloning site (MCS) was inserted into the vector. The MCS contains four frequently used restriction sites. Cloning a gene of interest in the new vector can now be a single step event. Furthermore, this new vector now allows one to carry out forced directional cloning of DNA fragments of interest.

MATERIALS AND METHODS

Plasmids and strains

The binary plant expression vector pPCV701 (9.6 kb) carries the mas P1,P2 promoters (Koncz *et al.*, 1987). The vector pPCV701 *luxFP1* (11.8 kb) is vector pPCV701 with the *fab2* luciferase gene cloned downstream of the P1 promoter (Escher, unpublished data). Competent cells were prepared with *E.coli* strains *DH5α* and *S17-1* as described previously (Inoue *et al.*, 1990).

Mobilization of binary plant expression vectors from *E. coli* into *A. tumefaciens*

Binary plant expression vectors were mobilized from *E. coli S17-1* into *A. tumefaciens* strain *GV3101 pMP 90RK* according to the method previously

described (Koncz and Schell, 1986). The transconjugants were screened *in vivo* using the Hamamatsu Argus 100 low-light video image analysis.

Isolation of binary plant expression vectors from *Agrobacterium*

Binary plant expression vectors were further analyzed by restriction enzyme mapping before use in plant transformation. The following procedures were used to isolate binary plant expression vectors from *Agrobacterium*: The YEB medium (1.5 ml) culture containing *Agrobacterium* strain *GV3101 pMP 90RK* and appropriate antibiotic selection was shaken (220 rpm) overnight at 29°C. The cells were pelleted for 3 minutes at 5,000 rpm in a microcentrifuge. Then, the pellet was resuspended in 300 µl 1 x TE buffer (pH 8.0) and centrifuged at 10,000 rpm for 1 minute in a microcentrifuge. The supernatant was removed and the cell pellet was resuspended in 100 µl 1 x TE buffer (pH 8.0). The cell suspension was incubated with 10 µl lysozyme (20 mg/ml) for 10 min at the room temperature, followed by alkaline lysis with sodium hydroxide (0.2 N NaOH + 1% SDS) and potassium acetate precipitation (3 M KoAc + 2 M HAc) for 10 min each on ice. After phenol-chloroform (1:1) extraction, the DNA was precipitated by 2.5 volume of 100% ethanol. The DNA pellet was resuspended in 40 µl ddH_2O from which approximately 5 µl was used for transformation of *E.coli* competent cells (Inoue *et al.*, 1990). Three ampicillin resistant *E. coli* colonies were randomly picked by sterile toothpicks and allowed to grow in 1.5 ml LB broth with appropriate antibiotic selection for 6 hours at 37°C with shaking (220 rpm). For isolation of binary plant expression vectors from *E. coli*, the procedure is the same as described above, but with 5 minute instead of 10 minute incubation for lysozyme, alkaline lysis and potassium acetate. The DNA, resuspended in 10 µl ddH_2O, was used for restriction enzyme mapping

according standard methods (Maniatis *et al.*, 1982). Based on low light image analysis and restriction enzyme mapping, the correct transconjugants were used to transform plant tissues.

Plant materials and transformation

Nicotiana tabacum cv. SR1 leaf disc explants (8-mm diameter) were harvested from young (4-5 leaf stage) plants grown on MS medium in sterile jars. After soaking for 15 minutes with log phase growth *A. tumefaciens* strain *GV3101 pMP 90RK* containing luciferase based binary vectors, the leaf discs were incubated on MS medium at 25°C, without antibiotic selection for two days in the dark. The leaf discs were transferred to MS shooter medium (MS medium + 0.1 mg/l of NAA + 0.5 mg/l of BAP) (Murashige and Skoog, 1962) containing 100 mg/l of kanamycin for selection of transformed plant cells and 400 mg/l claforan for preventing growth of *A. tumefaciens*. Kanamycin-resistant shoots appeared after three to four weeks and were excised from the leaf discs and rooted in jars containing MS medium without hormones but containing antibiotic selection. Roots developed two to three weeks later and the transformed plantlets were grown to maturity in soil in the greenhouse under a 12-hour-photoperiod as described previously (Jiang *et al.*, 1992).

Luciferase assays

A. tumefaciens transconjugants containing luciferase based binary plant transformation vectors were incubated overnight at 29°C and subjected to low-light image analysis for 2 minutes. Leaf, root and flower tissues from transgenic plants were screened for luciferase activity for 30 minutes. A detailed method of low-light image analysis with the Hamamatsu Argus 100

low-light video image analyzer has been described previously (Langridge *et al.*, 1991).

RESULTS AND DISCUSSION

Modification of the *fab2* gene

The mas promoters and the luciferase marker gene in vector pPCV701 *luxFP1* are useful tools for plant genetic engineering, because these dual (bidirectional) promoters allows coexpression of a bacterial luciferase with a gene of interest. Therefore, light emission in a transgenic plant is a strong indication of integration of a gene of choice into the genome of regenerated transgenic plants. However, a limitation of this vector is that the DNA fragment of the *fab2* marker gene contains sites which are identical to cloning sites in vector pPCV701 *luxFP1* (Fig. III-1a). The removal of these sites in the DNA fragment of the marker gene was attempted for cloning purposes. Bam HI and a Sal I sites are located 197 bp and 206 bp downstream from the stop codon of gene *fab2*, respectively (Escher *et al.*, 1989). In the first step the plasmid pPCV701 *luxFP1* was cut with restriction enzyme Bam HI to yield a 2.7 kb *fab2*-mas promoter DNA fragment which was further digested with Sal I. A 2.2 kb *fab2* Bam HI-Sal I DNA fragment was obtained and was ligated into the Sal I site downstream of the P1 promoter in vector pPCV701 (Fig. III-1b). After blunting the Bam HI site at the 3' end of *fab2* with Klenow fragment, the vector was recircularized by ligation. Competent cells of *E. coli DH5α* were transformed with the ligation mix, and recombinant plasmids were characterized using restriction analysis and agarose gel electrophoresis. A plasmid lacking the Bam HI site in the marker gene was identified and named pPCV701 *luxFM1* (Fig. III-1c).

Addition of a synthetic polylinker

In order to increase the number of cloning sites, a polylinker was added to the plant expression vector to enhance its versatility for cloning. For the selection of the proper restriction sites of the polylinker to be inserted into the vector, a group of frequently used restriction enzymes were tested on vector pPCV701 *luxFM1*. Based on the data obtained from restriction analysis, four commonly used restriction sites, Bam HI, Sma I, Sac I and Xba I were selected, and two complementary oligonucleotides were synthesized that contained sequences required by these restriction enzymes (Fig. III-2).

After phosphorylation and annealing, the double stranded synthetic MCS linker was isolated from a 15% polyacrylamide gel (Maniatis *et al.*, 1982). Vector pPCV701 *luxFM1* was digested with Bam HI, followed by alkaline phosphatase treatment. In order to keep only one Bam H1 site in the polylinker, one end of the MCS linker consisted of a Bgl II site, thereby preserving restoration of one of the Bam H1 site during ligation. Recombinant plasmids were characterized using restriction analysis and DNA sequencing (Sanger *et al.*, 1977). The desired plasmid was obtained and termed pPCV701 *luxFM2* (Fig. III-1d).

Luciferase assay

Although the Bam HI and Sal I sites are located downstream of the *fab2* gene, the above modifications could have affected the stability of the *fab2* messenger RNA, and have an effect on luciferase gene expression. Therefore, luciferase activities were compared both in *A. tumefaciens* and in transgenic tobacco plants containing the modified vector and the original vector. The *A. tumefaciens* strain *GV3101 pMP 90RK* containing the modified and original vectors were obtained from the conjugation experiment between *A.*

tumefaciens (*GV3101 pMP 90RK*) and *E. coli* (*S17-1*). Resulting transconjugants were subjected to low-light image analysis. Leaf disc explants from tobacco plant SR1 were transformed using *A. tumefaciens* strain *GV3101 pMP 90RK* containing the modified plant transformation vector, pPCV701 *luxFM2*, and the original vector, pPCV701 *luxFP1*. Kanamycin-resistant plants were regenerated and luciferase activity was compared. There were no differences in light emission observed both in *Agrobacteria* and transgenic plants containing the modified vector and the original vector (Fig. III-3) (data for transgenic plants not shown). The modified vector has been used successfully for cloning and expression of genes of interest in tobacco, tomato and *Arabidopsis* plants. For example, the vector has been used in creating an *in vivo* model system for monitoring specific induction of PAL promoter in *Arabidopsis* (Giacomin and Szalay, 1993), for studying ribozyme activity in tobacco (Ayre *et al.*, 1993), and for construction of a promoter test vector (see Chapter V).

In conclusion, a luciferase based binary plant expression vector pPCV701 *luxFP1* was modified and new vectors pPCV701 *luxFM1* and pPCV701 *luxFM2* were constructed with improved capabilities for expression of desired genes and the bacterial luciferase gene and for a study regulation of gene expression in higher plants.

BIBLIOGRAPHY

Ayre, B.G., Escher, A. and Szalay, A.A. (1993) Bioluminescence as a measure of ribozyme activity *in vivo*. Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds.) John Wiley & Sons. 109-112

Escher, A., O'Kane., D.J, Lee., J. and Szalay, A.A. (1989) Bacterial luciferase αβ fusion protein is fully active as a monomer and highly sensitive *in vivo* to elevated temperature. *Proc. Natl. Acad. Sci. USA* 86, 6528-32.

Gasser, C.S. and Fraley, R.T. (1989) Genetically engineering plants for crop improvement. *Science* 244, 1293-99.

Giacomin, L. and Szalay, A.A. (1993) The use of PAL-*lux* transcriptional fusion to monitor *in vivo* plant response to microbial invasion. Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds.) John Wiley & Sons. 207-211

Inoue, H., Norjima, H. and Okayama, H. (1990) High efficiency transformation of *E. coli* with plasmids. *Gene* 96, 23-28.

Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Identification of plant genes *in vivo* by tagging with T-DNA border-linked luciferase genes followed by inverse polymerase chain reaction amplification. *Plant. Mol. Biol. Rep.* 10, 345-61.

Koncz, C., Langridge, W.H.R., Olsson, O., Schell, J. and Szalay, A.A. (1990) Bacterial and firefly luciferase Genes in transgenic plants: advantages and disadvantages of a reporter gene. *Devel. Genetics* 11, 224-32.

Koncz, C., Olsson, O., Langridge, W.H.R., Schell, J. and Szalay, A.A. (1987) Expression and assembly of functional bacteria luciferase in Plants. *Proc. Natl. Acad. Sci. USA* 84, 131-35.

Koncz, C. and Schell, J. (1986) The promoter of the T1-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* 204, 383-96.

Langridge, W.H.R., Escher, A. and Szalay, A.A. (1991) Measurement of bacterial luciferase as a reporter enzyme *in vivo* in transformed bacteria, yeast and plant cells and in transgenic plants. *Technique* 3, 99-108.

Langridge, W.H.R., Fitzgerald, K.J., Koncz, C., Schell, J. and Szalay, A.A. (1989) Dual promoter of *Agrobacterium tumefaciens* mannopine synthase genes is regulated by plant growth hormones. *Proc. Natl. Acad. Sci. USA* 86, 3219-23.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant* . 15, 473-97.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbour Laboratory, Cold Spring Harbour, NY).

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463-67.

Schell, J. (1987) Transgenic plants as tools to study the molecular organization of plant genes. *Science* 237, 1176-82.

Velten, J., Velten, L., Hain, R. and Schell, J. (1984) Isolation of a dual plant promoter fragment from the Ti plasmid of *Agrobacterium tumefaciens*. *EMBO J.* 3, 2723-30.

FIGURES AND LEGENDS

Fig. III-1. Diagram depicting the strategy employed in the construction of vectors, pPCV701 *luxFM1* and pPCV701 *luxFM2*. A luciferase based binary plant transformation vector pPCV701 *luxFP1* (**a**) was digested with Bam HI and Sal I respectively and the DNA fragment containing the *fab2* gene with a Sal I site at the 5' end and a Bam HI site at the 3' end was ligated to a vector, pPCV701 (**b**), on the Sal I site. This ligation resulted a linear vector with incompatible ends (Bam HI and Sal I). After blunt ending, the vector was recircularized by ligation and named pPCV701 *luxFM1* (**c**). A MCS (multiple cloning site) was added into the Bam HI site in the vector, pPCV701 *luxFM1* and the new vector was named pPCV701 *luxFM2* (**d**). The T-DNA between right and left borders (RB and LB) indicates the transferred DNA which is stably integrated into the plant genome. The arrows indicate the direction of transcription; P_{NOS} is nopaline synthase promoter, Pg5 is the promoter of T_L-DNA gene 5; g4pA, OcspA, and g7pA are polyadenylation sequences derived from T_L-DNA gene 4, the octopine synthase gene, and gene 7; NPTII is neomycin phosphotransferase gene; ori_T and ori_V are replication and conjugational-transfer origin sequences derived from plasmid pRK2; oripBR is replication of origin of pBR322; Ap^r is the β-lactamase gene from *E. coli*.

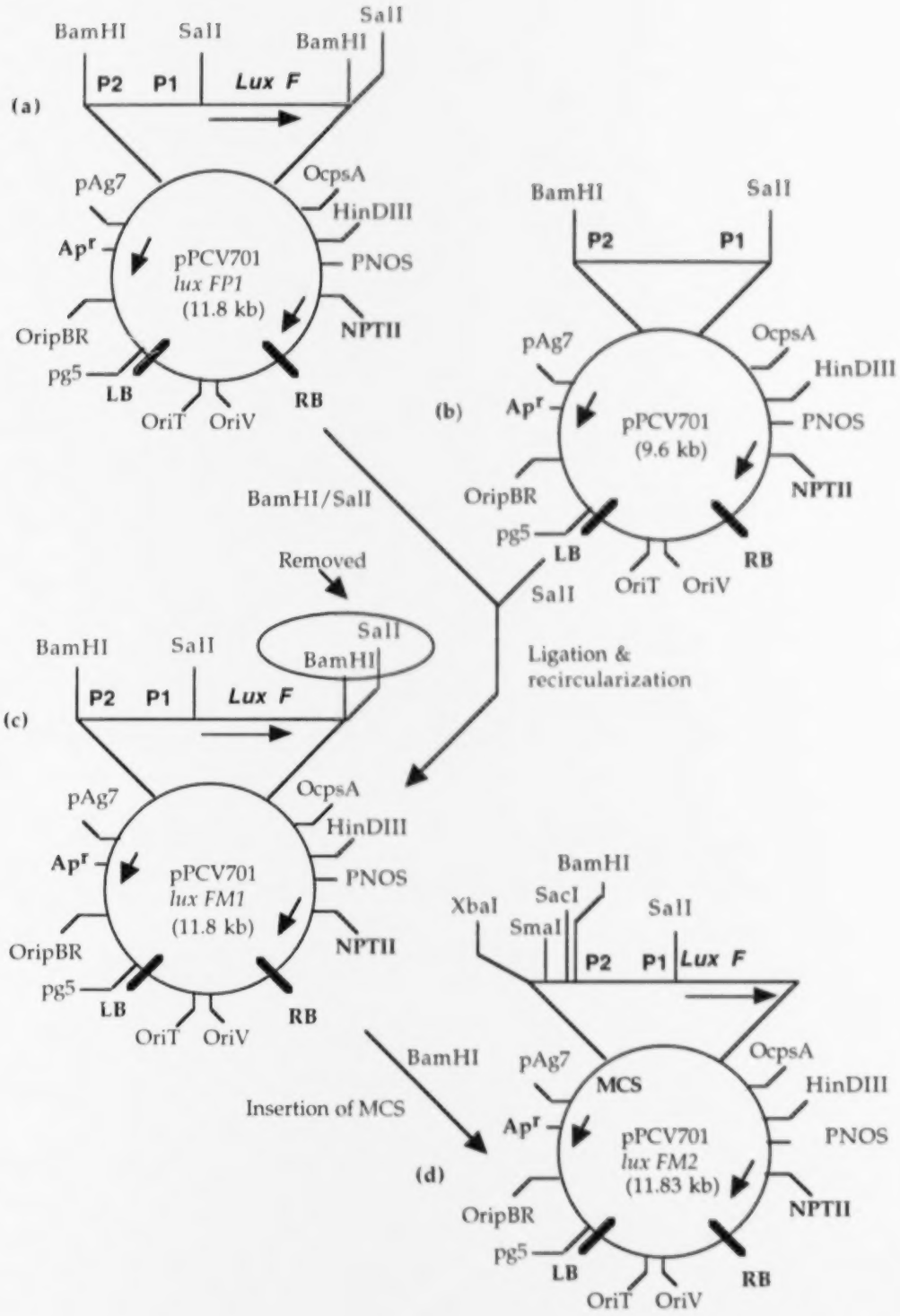

(a)
BamHI
SalI
BamHI
SalI
P2
P1
Lux F
pAg7
OcpsA
HinDIII
Ap^r
pPCV701
lux FP1
(11.8 kb)
PNOS
NPTII
OripBR
pg5
LB
RB
OriT
OriV
(b)
BamHI
SalI
P2
P1
pAg7
OcpsA
HinDIII
Ap^r
pPCV701
(9.6 kb)
PNOS
NPTII
OripBR
pg5
LB
RB
OriT
OriV
BamHI/SalI
Removed
SalI
Ligation &
recircularization
(c)
BamHI
SalI
BamHI
P2
P1
Lux F
pAg7
OcpsA
HinDIII
Ap^r
pPCV701
lux FM1
(11.8 kb)
PNOS
NPTII
OripBR
pg5
LB
RB
OriT
OriV
BamHI
XbaI
SacI
SmaI
SalI
P2
P1
Lux F
BamHI
Insertion of MCS
pAg7
MCS
OcpsA
HinDIII
Ap^r
pPCV701
lux FM2
(11.83 kb)
PNOS
NPTII
(d)
OripBR
pg5
LB
RB
OriT
OriV

Fig. III-2. DNA sequence of the synthetic multiple cloning site (MCS). The sequence is based on the data obtained from restriction analysis. Four commonly used restriction sites such as Bam HI, Sma I, Sac I and Xba I were selected and two partially complementary oligonucleotides containing DNA sequences required by these restriction enzymes were synthesized. After phosphorylation and annealing, the double stranded synthetic MCS linker was purified from a 15% polyacrylamide gel. The Bgl II site was designed so as to preserve only one Bam HI site in the vector. The MCS linker was ligated to the Bam HI site at the downstream of the P2 promoter in the vector pPCV701 *luxFM1* and formed a new vector pPCV701 *luxFM2*.

Bam HI Sma I Sac I Xba I

5'-**GATCCCCGGGAGCTCTCTAGA**-3'

3'-**GGGCCCTCGAGAGATCTCTAG**-5'

Bgl II

The material involved in page 91 has been removed because of copyright restrictions. This page contains a legend of photographs which has been published. Jiang, C., Escher, A. and Szalay, A.A. (1993) Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds.) John Wiley & Sons. 130.

The material involved in page 92 has been removed because of copyright restrictions. This page contains photographs which have been published. Jiang, C., Escher, A. and Szalay, A.A. (1993) Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds.) John Wiley & Sons. 130.

Chapter IV

Identification of Activation of Plant Genes in vivo by Tagging with a T-DNA Border-Linked Luciferase Gene

1. INTRODUCTION

Three methods are currently available for isolation of plant genes whose products have not yet been characterized. The first and most frequently used method is based on chromosome walking with the help of cosmid or yeast artificial chromosome cloning systems (Olson *et al.*, 1986; Green and Olson, 1990). The gene is localized on a restriction fragment containing an RFLP marker (Chang *et al.*, 1988). This method is laborious and can be impeded by large stretches of repetitive DNA (Brenda *et al.*, 1992; Sun *et al.*, 1992). The second approach is genomic subtraction. This procedure depends on the presence of a characterized large deletion (larger than 5 kb) in the mutant plant genome when compared with the wild-type genome (Straus and Ausubel, 1990; Sun *et al.*, 1992). The third method is based on the T-DNA tagging. This method relies on the T-DNA border-based integration of a known foreign DNA sequence into a plant genome (Koncz *et al.*, 1989; Feldmann *et al.*, 1989). The T-DNA border sequence serves as a mutator via interruption of the plant genome and as a signal for isolation of the interrupted sequences by insertion of sequences linked to the flanking genomic DNA.

T-DNA tagging is an efficient method for creation of mutants and isolation of mutant genes from plants (Koncz *et al.*, 1990; Feldmann, 1991). Since integration of T-DNA into the plant genome may result in rearrangement of newly integrated DNA, plasmid rescue methods are not always successful for isolation of tagged DNA fragments as deletions may occur, resulting in removal of part of a plasmid replicon, or the antibiotic selection marker gene. To overcome these disadvantages, T-DNA tagging was combined with inverse polymerase chain reaction amplification of target DNA sequence (Ochman *et al.*, 1988). In addition, a bacterial luciferase

marker gene system (Olsson *et al.*, 1989; Langridge *et al.*, 1991) was added so that a mutant generated by T-DNA insertion can be verified *in vivo* by a simple method of screening transformants by low-light video image analysis. By combining the T-DNA tagging method with a promoterless luciferase gene fusion followed by iPCR amplification, a developmentally regulated promoter element was isolated from tobacco plants. In addition to plant gene identification, the luciferase marker gene system can be used directly for isolation of promoter and enhancer elements, as well as gene identification in organisms in which gene delivery systems are available, such as yeast and *Drosophila*.

2. MATERIALS AND METHODS

Bacterial strains and DNA manipulation

Competent cells were prepared with *E. coli* strains *DH5α* and *S17-1* as described previously (Inoue *et al.*, 1990). A plant insertional vector, pPCVG *luxA&B*, was mobilized from *S17-1* into *A. tumefaciens* strain *GV3101 pMP 90RK* according to the method described previously (Koncz and Schell, 1986). An appropriate transconjugant was selected according to the method described in Chapter III of this thesis and used to transform tobacco plants. DNA manipulation involved in the vector construction was carried out by using standard protocols (Sambrook *et al.*, 1989).

Agrobacterium-mediated transformation of tobacco leaf explants and screening of regenerated transgenic plants

Nicotiana tabacum cv. SR1 leaf disc explants (8-mm diameter) were excised from sterile plants grown in jars and transformed by a 15-minute soaking of the leaf discs in a 9-cm culture dish filled with a log phase growth of *A. tumefaciens* strain *GV3101 pMP 90RK* containing the plant insertional

vector pPCVG *luxA&B* (see Fig. IV-1). The infected leaf disc explants were incubated on MS medium (Murashige and Skoog, 1962) supplemented with 3% sucrose in 9-cm plastic culture dishes at 25°C, without antibiotic selection, for 48 hours in the dark. The leaf explants then were transferred to MS shooter medium (MS medium supplemented with 3% sucrose + 0.1 mg/l of naphthalene acetic acid (NAA) + 0.5 mg/l of benzylamino purine (BAP) in 9-cm plastic culture dishes containing 400 mg/l of claforan and 20 mg/l of hygromycin and incubated at 25°C under a 18-hour photoperiod. Hygromycin-resistant shoots, which appeared after three to four weeks, were excised from the leaf discs and rooted in jars containing MS medium without hormones but containing hygromycin and claforan at the levels previously described (Chapter II). Roots developed in two to three weeks and the plantlets were grown to maturity in soil in the greenhouse at 18°C under a 12-hour photoperiod.

***In-vivo* detection of insertional mutant plants**

A rapid and sensitive *in vivo* assay method for the detection of luciferase gene expression was developed to detect the time and precise location of gene expression in transgenic plant tissues as follows. Low-light video image analysis of intact plantlets, plant leaves, flowers, roots and stem tissues was performed with an Argus 100, low-light Video Image VIM Camera (Hamamatsu Photonics,K.K., Hamamatsu City, 431-32, Japan). Young hygromycin-resistant tobacco plants were screened for luciferase marker gene expression by placing the entire plant into a large (19-cm) plastic culture dish containing filter paper impregnated with a 10% aqueous solution of the luciferase substrate decanal (Sigma). The closed culture dish was placed into imaging chamber of a Hamamatsu Argus 100 low-light video image analyzer. The chamber door was closed to exclude external light and after waiting for

two minutes for the extinction of fluorescence from chlorophyll, bioluminescence from the plant tissues was collected over a ten to thirty minute collection interval. The bioluminescence image was superimposed upon a video image of the plant to determine the exact location of the bioluminescence in the plant tissue. After analysis, the tested plants were returned to soil for further growth and reanalysis at a later stage of development. Individual flowers, leaf, stem and root tissues were excised from mature plants for low-light video image analysis exactly as described for young plantlets. Alternatively, light emission measurements were made on homogenates of plant leaves, roots, stems and flowers with 500 µl of lux buffer (50 mM $NaHPO_4.7H_2O$, pH 7.0; 50 mM S-EtOH; 0.4M sucrose; 0.1% BSA) in a luminometer (Langridge *et al.*, 1991). Luminometry of plant tissue samples was performed with a Turner Model 2e Luminometer (Turner Designs, Sunnyvale CA., 49086).

Isolation of genomic DNA from transgenic plant

For DNA isolation, seeds were harvested from the transgenic tobacco plant of interest and germinated and grown on MS medium minus hormones, containing 20 mg/l hygromycin. The seedlings were grown to the four- to five-leaf stage in the greenhouse at 18°C under 12-hour photoperiod. Prior to harvest of the leaves, plants were placed in the dark at room temperature for three days to reduce the level of polysaccharides. Young fully expanded leaves were excised from the plant and total DNA was isolated from leaf tissues by a CTAB extraction method (Draper *et al.*,1989), with the exception that 2 x CTAB concentration was used and the incubation temperature was increased to 60°C. This method resulted in high quality DNA (larger than 23 kb in size).

Identification of tobacco DNA segments linked to the luciferase gene by Southern hybridization.

To isolate the putative promoter sequence located upstream of the luciferase gene, transgenic tobacco plant DNA was cleaved with several restriction endonucleases. Tobacco genomic DNA (10 µg), was digested with the restriction endonuclease enzymes including Bam HI, EcoR I Hind III, and Xba I, and the DNA fragments were separated by electrophoresis in 0.5 x TBE buffer (0.089 M Tris-borate, 0.089 M boric acid, 0.002 M EDTA) on an 0.8% agarose gel (15 x 17 cm) at 40 V overnight (Sambrook *et al.*, 1989). The DNA was transferred to a Hybond-N nylon membrane (Amersham, Life Sciences Co.) by a vacuum-blotting procedure (Hoffer Scientific Instrument Co.). The DNA was then fixed on the membrane by UV crosslinking using a UV Crossliker (Stratagene). The *lux A* DNA probes were labeled using a random prime labeling kit (Boehringer Mannheim) according to the included protocol with the exception that the random primer labelling procedure was carried out for at least 4 hours at room temperature instead of 1 hour at 37°C. Prehybridization and hybridization of the genomic DNA fragments to a ^{32}P labelled *luxA* gene probe were carried out for 3 and 20 hours, respectively, at 42°C in pyrex hybridization tubes in a Robbins hybridization incubator (Model 310, Robbins Scientific Co., Sunnyvale CA.) with a mixture of 50% formamide; 6 X SSC; 50 mM phosphate buffer, pH 7.0; 5 X Denhardts; 1% SDS; 150 µg/ml salmon sperm DNA (heat denatured). After hybridization, the membrane was washed once in 2 X SSC, 0.1% SDS at room temperature and two times at 42°C and followed by washing once in 0.2 X SSC, 0.1% SDS at 42°C and once at 68°C. The dried membrane was exposed to Kodak XAR-5 film (Kodak X-Omat XK-1) in a photo-cassette equipped with intensifying screen in the dark for two to three days at -70°C.

Isolation and self-ligation of *lux*-containing genomic DNA fragments from transgenic plants for inverse PCR

Isolation and amplification of the DNA fragment flanking the insertion can be accomplished in a single step based on the hybridization of appropriately designed specific oligonucleotide primers to the DNA sequence adjacent to the target followed by iPCR amplification of the target sequence.

For iPCR (Ochman *et al.*, 1988), total genomic DNA (30 µg) from the transgenic plant 12-1 was digested with 100 U of Xba I in 600 µl total volume for 4 hrs. at 37°C. When the digestion was completed, the enzyme was heat-inactivated at 70°C for 10 minutes and extracted by phenol/chloroform (1:1) followed by precipitation of the DNA with 100% ethanol. The pellet was resuspended in 100 µl of ddH_2O and loaded on an 0.8% agarose gel (15 X 17 cm). Electrophoresis was carried out 40 V overnight in TBE buffer. The DNA bands corresponding to the Southern hybridization data were collected with dialysis tubing and purified by Elutip-D column. The ligation mixture (500 µl) containing 1 µg of DNA fragments, 20 U of T4 ligase 0.5 µM of ATP and 1 X T4 ligation buffer was incubated at 16°C overnight (Boehringer Mannheim). The T4 DNA ligase in the preparation was then heat-inactivated by incubation at 75°C for 10 min. The mixture was extracted once with phenol/chloroform (1:1) and precipitated by addition of 0.1 vol of 3.0 M of sodium acetate and 2.5 vol of 100% ice-cold ethanol. The DNA template was air-dried and resuspended in 10 µl 1 X TE buffer (pH 8.0).

Two oligonucleotide primers whose sequence is complementary to the 3' and the 5' end of the coding region of the *lux A* gene (Cohn *et al.*, 1985) were synthesized on a model 380 B DNA Synthesizer (Applied Biosystems, California). Unlike normal PCR primers, the orientation of these iPCR primers is in the outward direction, exiting from the *luxA* gene. The base

sequences of the iPCR primers exiting upstream and downstream of the *Lux A* gene were 5'-CTGATAAGTGAGAAGGAAGTTTCC-3' and 5'-GCTATTCCAGTCTGATGTATGCC-3', respectively.

Conditions for inverse PCR

Polymerase chain reaction amplification of plant DNA sequences was performed in a Thermal Cycler, Model N 801-0150 (Perkin Elmer-Cetus Corp., Norwalk, CT, USA). In a 500 µl microcentrifuge tube, the following reagents were mixed in order: denatured template DNA (1 µg); 10 X reaction buffer; 10 mM dNTPs; and 100 µM MgCl (Boehringer Mannheim). The DNA Thermal Cycler was heated to 70^{o}C, and 8.0 µM each of the iPCR primers and 2 units of Taq DNA polymerase were added to the reaction tube. The tube contents were mixed by vortexing, centrifuged briefly and a layer of 30 µl of paraffin oil was placed on top. Ten pmol primer (1/10 of the total primer) was added for the first 10 cycles to avoid nonspecific DNA amplification. When the iPCR reached the 10th cycle, the remaining primers were added to the reaction mixture. When the iPCR reached 20 cycles, an additional 2 units of Taq polymerase was added to the reaction mixture to compensate for losses in polymerase activity. The iPCR conditions were 40 cycles at 94^{o}C for 1 min, 61^{o}C for 2 min, and 72^{o}C for 3 min. After the last cycle, the PCR mixture was incubated at 72^{o}C for 10 min to complete the amplification, (e.g. utilization of unreacted primers etc.). The reaction mixture was removed from the thermal cycler and a 10 µl sample was subjected to electrophoresis on a 0.8% agarose mini-gel to determine the extent of amplification and the number and quality of the DNA products. The tube containing the amplification reaction was stored at -20^{o}C for one hour until the PCR reaction mixture was frozen. The oil layer on the top of the reaction mixture could then be removed easily with a micropipette without losing any of the reaction mixture.

RESULTS AND DISCUSSION

Construction of a specialized plant insertional vector and experimental strategy

A specialized plant insertional vector used for the present study, pPCVG *luxA&B*, was constructed by using a promoterless *lux A* gene form a marine bacterium *Vibrio harveyi* (Fig. IV-1). The vector has several features: The inclusion of the promoterless luciferase marker gene linked to the right border of the T-DNA in a plant insertional vector provides a tool for simple, *in vivo* screening for insertional mutants in transgenic plants. The CaMV 35S RNA promoter was added to the insertional vector. The 35S RNA promoter not only serves as promoter for the *luxB* gene, but also offers enhancement for the *luxA* gene, as the enhancer region of the 35S RNA promoter can function in both orientations. Consequently, it was possible to identify insertional mutants with organ or stage specific gene expression patterns by enhancement of the luciferase gene expression *in vivo* as described in the following section. Inclusion of *hpt* (hygromycin phosphotransferase) gene in the vector allows selection of transgenic plants base on a high level of hygromycin resistance. The strength of this approach is that the activation of the luciferase gene during development in a transformed plant may reflect the pattern of native gene expression. Therefore, identification of insertional mutants can be based on marker gene expression rather than on phenotype. A schematic representation of the hypothesis of insertional activation of the promoterless luciferase marker gene in transgenic plants is illustrated in Fig. IV-2. If a T-DNA carrying a promoterless luciferase gene is integrated in just downstream of a plant promoter, the *lux* gene will be activated by plant promoter signals resulting

expression of luciferase gene in a transgenic plant which can be detected *in vivo* by low-light video image analysis (Fig. IV-2b). These type of insertional mutants can not be identified by nonspecialized T-DNA tagging methods. Therefore, inclusion of the promoterless luciferase gene in the vector not only offers an simple *in vivo* screening method, but also greatly enhances T-DNA tagging efficiency. As Fig. IV-2c illustrated, when T-DNA is integrated in the middle of a plant gene, the luciferase activity may or may not be detected depending on particular situation of transcription. However, these type of insertional mutants may be identified based on homozygous mutant phenotype if gene products are essential for plant development.

Selection of insertional mutants based on the luciferase gene expression *in vivo* during plant development

In *Agrobacterium-mediated* transformation experiments containing the plant insertional vector, pPCVG *luxA&B*, a large number of hygromycin-resistant transformed plants were recovered and regenerated. During the growth and development of these transformed plants, systematic luciferase assays were carried out *in vivo* (at early stages) and *in vitro* (at late stages) to monitor gene activation in a native environment. In screening experiments, *lux* gene expression in 45 (17.8%) different transformants was obtained from a total of 253 hygromycin-resistant tobacco plants. Table IV-1 summarizes the results from the assays. Based on these results, insertional mutants were selected. Among 45 luciferase positive transgenic plants screened, 42% showed constitutive luciferase gene expression patterns. However, in several transgenic plants luciferase activity was detected in predominantly one organ (11%). The remaining transgenic plants (46%) expressed luciferase in more than one organ but not constitutively in all plant organs. Fig. IV-3 shows the specific localization of luciferase gene expression in some of those transgenic

tobacco plants. In addition, two independent transgenic plants selected demonstrated constitutive luciferase activity in different parts of the plants (Chart IV-1). In the LTM-1 plants, luciferase activity increased from the top leaf to the bottom leaf as well as from the top stem to the bottom. However, the luciferase activity in the second plant (LTM-2) decreased from the top to the bottom in both leaves and stems. Taken together, these results prove that mutant selection based on luciferase gene expression can be carried out on the original transformed plants rather than in the second generation of mutant plants as previously reported (Feldmann, 1991). Therefore, this approach greatly enhances the efficiency for the mutant selection. More importantly, this approach provides a convenient tool to identify a mutant with a temporal gene activation pattern that may not be identified by other methods (Walden *et al.*, 1991).

T-DNA insertional mutagenesis has been shown to be a powerful method for isolation of genes and regulatory sequences from plants (Koncz *et al.*, 1990; Yanofsky *et al.*, 1990; Feldmann, 1991; Kertbundit *et al.*, 1991). However, the major concern in previous T-DNA tagging methods is their low mutation frequency (Walden *et al.*, 1991). With the promoterless bacterial luciferase gene linked to the right border of the T-DNA, a mutant detection system established in this study permitted possible selection of mutants based on the activation of light emission rather than by phenotypic alteration. This feature allows us to apply T-DNA tagging in polyploid species. Further, the bacterial luciferase reporter gene is known to offer reproducible *in vitro* and *in vivo* assay systems which is useful for detection of spatial and temporal expression of gene activation in real time during the life cycle of the transgenic plant. This is a unique application of the T-DNA-bacterial luciferase gene tagging system in comparison to other reporter gene systems

such as GUS (Jefferson *et al.*, 1987; Walden *et al.*, 1990; Kertbundit *et al.*,1991) which require destructive analysis of plant tissues to determine gene expression levels.

Localization of the T-DNA insert by Southern hybridization

One insertional mutant (plant 12-1) expressed luciferase gene predominantly in flowers was studied in detail. For segregation analysis of the 12-1 plant, self-pollination seeds were harvested from the primary transgenic plant and 200 seeds per plate were placed on MS medium containing 20 mg/l of hygromycin. The results showed a 3 : 1 ratio of hygromycin resistant vs. sensitive plants. For a detailed analysis, genomic DNA was isolated from young leaf tissue of plant 12-1 and digested to completion with restriction enzymes, Xba I, Hind III / Xba I, EcoR I and Hind III (Fig. IV-4). DNA hybridization with a radio-labelled DNA fragment of the *lux A* gene as a probe showed that 12-1 transgenic plant contains a single copy of the T-DNA insert. Several hybridization bands were detected including a single 2.5 kb Xba I band which might contain a 1.1 kb of flanking plant DNA located upstream of the promoterless lux A gene. One 2.3 kb Xba I / Hind III band was also detected. Because there is no Hind III site inside of the vector, the cutting must take place in the flanking plant DNA region.

Amplification of plant genomic DNA flanking the T-DNA insertion by inverse polymerase chain reaction

The approximate location of the *lux* gene containing a tobacco genomic DNA fragment was identified by Southern hybridization experiments (see Fig. IV-4). The target sequence can be rescued in *E. coli* (Koncz *et al*, 1989), however, deletions in the plasmid replicon or the antibiotic resistance marker gene can make this approach difficult or impossible (as was the case in my earlier attempts to clone the *lux* containing fragment). To circumvent these

cloning difficulties and to simultaneously amplify the *lux* containing DNA fragment, the iPCR method was used to facilitate amplification of unknown DNA sequences adjacent to a known DNA sequence (Ochman *et al.* 1988). The schematic procedure of inverse PCR is illustrated in Fig. IV-5.

Based on the results from Southern hybridization and sequence data from the *Lux A* gene, two PCR primers were designed to amplify the flanking plant DNA sequence. The primers specific to the ends of *lux A* gene and reading 'out' into the flanking plant DNA were annealed to the DNA template to amplify the flanking DNA sequence. A 1.4 kb DNA fragment of PCR products was obtained in a reaction containing a DNA template from the plant 12-1 with ligation, but no band was found both in a reaction containing the same template without ligation and a template from a nontransgenic plant (SR1) (Fig. IV-6). The size of the amplified fragment matched the size expected based on the plasmid map and the Southern hybridization data.

In vitro DNA sequence amplification by the polymerase chain reaction (PCR) has become one of the most powerful techniques for the analysis of gene expression available today (Saiki *et al.*, 1985). Since introduction of the inverse PCR method (Ochman *et al.*, 1988), application of iPCR has facilitated the isolation of sequences upstream or downstream of a known DNA sequence. Using this method, a flanking plant DNA sequence has been isolated (see Fig. IV-6).

By integration of T-DNA sequences into the plant genome, the T-DNA tagging method generates mutants by interrupting gene function (Walden *et al.*, 1991). The results of this study show that inclusion of the promoterless luciferase gene linked to the right T-DNA border in the plant insertional vector pPCVG *luxA&B* provides a tool for simple, *in vivo* screening for upstream control regions of developmentally regulated plant genes. Coupled

with specifically designed primers, inverse PCR methods provide a fast procedure for isolation and *in vitro* amplification of DNA sequences flanking the insertion. The combination of T-DNA tagging with the luciferase gene fusion, coupled with iPCR amplification of DNA sequences flanking the insertion, provides a powerful tool for isolation of plant genes and their regulatory sequences. The promoterless luciferase marker gene system linked to a transposable element can be used to isolate genes from other plants, insects and animal species in which transposon tagging or other gene delivery systems are available.

Additional questions relating to chapter IV

Is T-DNA integration random or non-random?

The data generated from this experiment provided additional evidence to support that T-DNA integration is a non-random event. In an analysis of 253 transformed tobacco plants, 45 or 1/5 of the population expressed the introduced promoterless luciferase gene. In other word, 20% of the transformed plants contained the T-DNA insert integrated within transcribed regions. The observed frequency of *lux* gene activation in this experiment is similar to that previously reported in both *Arabidopsis* and tobacco using a promoterless *npt II* gene (Koncz *et al.*, 1989), and is also similar to that found by Herman *et al* (1990). Much higher frequency (78% of all transformants) of insertional activation of promoterless *gus* gene has also been reported (Topping *et al.*, 1991). This frequency is higher than would be expected based on random integration of T-DNA insertion. First, it was estimated that tobacco contain approximately 60.000 structural genes in its 2.4 X 10^6 kb size of genome (Kamalay and Goldberg, 1980 and 1984). If an average gene in tobacco is 5 kb in size, the DNA content of structural genes is 3 X 10^5 kb accounting for

only 12.5% of the total DNA content. Thus, if T-DNA integration is random, the maximum frequency expected for *lux* gene activation should be around 12.5%. Second, the frequency observed may reflect only a portion of the T-DNA integration in the transcribed regions. As T-DNA integration within a transcriptional unit is non-polar (Herman *et al.*, 1990), it is estimated that only 50% of inserts within transcriptionally active DNA will be detected by activation of the promoterless luciferase genes with respect to endogenous promoters, insertion position and orientation of the T-DNA. Therefore, the frequency should be doubled in order to more accurately reflect the T-DNA integration within transcribed regions. Furthermore, the assay system used in this experiment depends on translation of a promoterless luciferase gene into a functional enzyme. Therefore, factors affecting protein translation may also influence the possibility of detecting activation of the *lux* gene. For example, possible translational fusion proteins generated from the *lux* gene and coding sequences of native genes may not be detected by luciferase assay. Taken together, the frequency observed on *lux* activation in this experiment suggests that the T-DNA integration into tobacco genome is not random.

The rationale for choosing amphidiploid tobacco

Agrobacterium tumefaciens has been successfully used to transform numerous plant species (Gasser and Fraley, 1989). However, highly efficient transformation protocols exist for only a limited number of plant species. Tobacco and *Arabidopsis* are two of these. As was discussed in chapter II, section 5, the different genome size between these two species does not affect the frequency of gene fusions (Koncz *et al.*, 1989). Although, it is difficult to identify phenotypic mutants from tobacco because of it amphidiploid status, there should be no difference between these two species when the selection is based on the activation of T-DNA-reporter genes (Koncz *et al.*, 1989). As this

experiment was designed to investigate the activation of the promoterless *lux* gene *in vivo* upon T-DNA integration, tobacco is an ideal candidate for this purpose. The additional advantage of working with tobacco is that the overall patterns of gene expression have already been intensely studied throughout the life cycle of the plant (Kamalay and Goldberg, 1980 and 1984). This will allow more meaningful comparisons between the the frequency of gene activation and patterns of expression to be made.

BIBILIOGRAPHY

Brenda, W.S., Susan, H. and Goodman, H.M.G. (1992) Effects of ionizing radiation on a plant genome: analysis of two *Arabidopsis transparent testa* mutations. *Plant Cell* 4, 333-47.

Chang, C., Bowman, J.L., Dejohn, A.W., Lander, E.S. and Meyerowitz, E.M. (1988) Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 85, 68-6860.

Cohn, D.H., Mileham, A.J., Simon, M.I., Nealson, K.H., Rausch, S.K., Bonam, D. and Baldwin, T. (1985) Nucleotide sequence of the *luxA* gene of *Vibrio harveyi* and the complete amino acid sequence of the alpha subunit of bacterial luciferase. *J. Biol. Chem.* 260, 6139 - 64.

Collins, F.S. and Weissman, S.M. (1984) Directional cloning of DNA fragments at a large distance from an initial probe: A circularization method. *Proc. Natl. Acad. Sci. USA* 81, 6812-16.

Draper, J., Scott, R. and Armitage, P. (1989) Plant Genetic Transformation and Gene Expression. A laboratory manual. Blackwell Scientific Publications, Oxford. 199-208.

Feldmann, K. A. (1991) T-DNA insertion in *Arabidopsis*: mutational spectrum. *Plant J.* 1, 71-82.

Feldmann, K.A., Marks, M.D., Christianson, M.L. and Quatrano, R.S. (1989) A dwarf mutant of *Arabidopsis* generated by T-DNA insertion mutagenesis. *Science* 243, 1351-54.

Gasser, C.S. and Fraley, R.T. (1989) Genetically engineering plants for crop improvement. *Science* 244, 1293-99.

Green, E.D. and Olson, M.V. (1990) Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: a model for human genome mapping. *Science* 250, 94-98.

Herman, L., Jacobs, A., Van Montagu, M. and Depicker, A. (1990) Plant chromosome/marker gene fusion assay for study of normal and truncated T-DNA integration events. *Mol. Gen. Genet.* 224, 248-56.

Inoue, H., Norjima, H. and Okayama, H. (1990) High efficiency transformation of *E. coli* with plasmids. *Gene* 96, 23-28.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M. (1987) GUS fusion: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6, 3901-7.

Kamalay, J.C. and Goldberg, R.B. (1980) Regulation of structural gene expression in tobacco. *Cell* 19, 935-46.

Kamalay, J.C. and Goldberg, R.B. (1984) Organ-specific nuclear RNAs in tobacco. *Proc. Natl. Acad. Sci. USA* 81, 2081-85.

Kertbundit, S., De Greve, H., DeBoeck, F., Van Montagu, M. and Hernalsteens, J.P. (1991) *In vivo* random β-glucuronidase gene fusion in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 88, 5212-16.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Korber, H., Redei, G.P. and Schell, J. (1989) High-frequency T-DNA-mediated gene tagging in plants. *Proc. Natl. Acad. Sci. USA* 86, 8467-71.

Koncz, C., Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Reiss, B., Redei, G.P. and Schell, J. (1990) Isolation of a gene encoding a novel chloroplast protein by T-DNA tagging in *Arabidopsis thaliana*. *EMBO J.* 9, 1337-46.

Koncz, C. and Schell, J. (1986) The promoter of the T1-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* 204, 383-96.

Langridge, W.H.R., Escher, A.P. and Szalay, A.A. (1991) Measurement of bacterial luciferase as a reporter enzyme *in vivo* in transformed bacteria, yeast and plant cells and in transgenic plants. *Technique* 3, 99-108.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* 15, 473-97.

Ochman, H., Gerber, A.S. and Hartl, L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* 120, 621-23.

Olson, M.V., Dutchik, J., Graham, M., Broder, G., Helms, C., Frank, M., MacCollin, M., Scheinman, R. and Frank, T. (1986) Random-clone strategy for genomic restriction mapping in yeast. *Proc. Natl. Acad. Sci. USA* 83, 7821-25.

Olsson, O., Escher, A.P., Sandberg, G., Schell, J., Koncz, C. and Szalay, A.A. (1989) Engineering of monomeric bacterial luciferases by fusion of *luxA* and *luxB* genes in *Vibrio harveyi*. *Gene* 81, 335-47.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985) Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350-54.

Sambrook, J., Maniatis, T. and Fritsch, E.F. (1989) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sanger, F., Niklen, S. and Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463-67.

Straus, D. and Ausubel, F.M. (1990) Genomic subtraction for cloning DNA corresponding to deletion mutants. *Proc. Natl. Acad. Sci. USA* 87, 1889-93.

Sun, T.P., Goodman, H.M. and Ausubel, E.M. (1992) Cloning the *Arabidopsis GA1* locus by genomic subtraction. *Plant Cell* 4, 119-28.

Walden, R., Hayashit, H and Schell. J. (1991) T-DNA as a gene tag. *Plant J.* 1, 281-88.

Walden, R., Koncz, C. and Schell, J. (1990) The use of gene vectors in plant molecular biology. *Methods Mol. Cell Biol.* 1, 175-94.

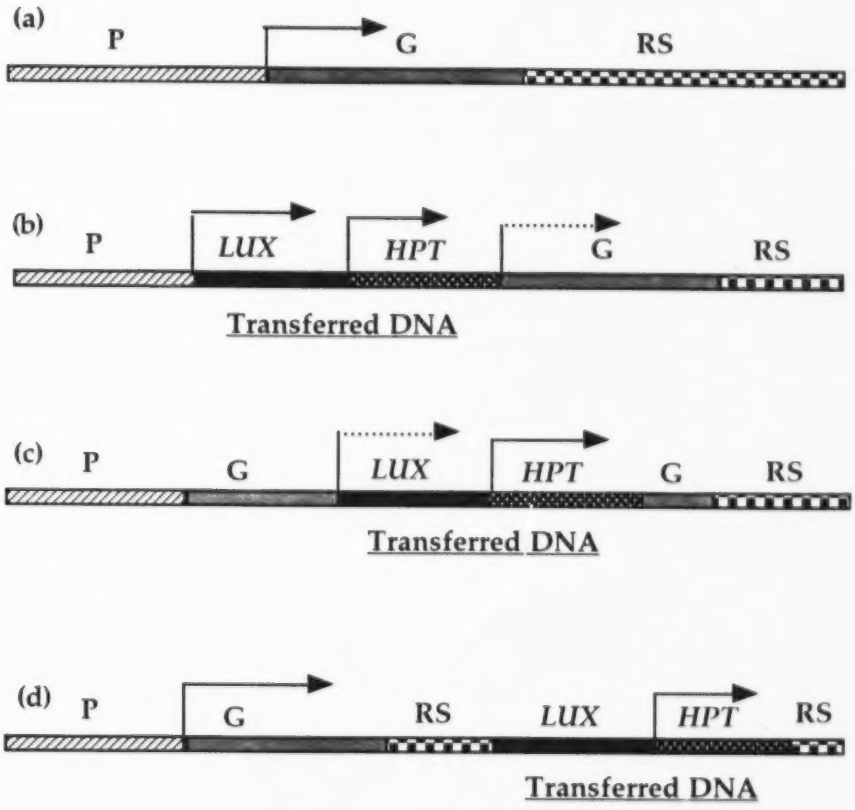
Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A. and Meyerowitz, E.M. (1990) The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature* 346, 35-39.

FIGURES AND LEGENDS

The material involved in page 114 has been removed because of copyright restrictions. This page contains a legend of a circular map of the plant insertional vector pPCVG *luxA&B* . Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Plant Mol. Biol. Rep. 10, 348.

The material involved in page 115 has been removed because of copyright restrictions. This page contains a circular map of the plant insertional vector pPCVG *luxA&B* . Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Plant Mol. Biol. Rep. 10, 348.

Fig. IV-2. Schematic representation of insertional activation of a promoterless bacterial luciferase marker gene in the plant genome. **Fig. IV-2a** represents a plant gene and its promoter sequences before the integration of a T-DNA (transferred DNA). T-DNA with a promoterless luciferase marker gene integrates immediately downstream of a plant promoter and, therefore, luciferase messenger RNA is made and luciferase activity can detected in transgenic plants (**Fig. IV-2b**). T-DNA inserts in the inside of a gene and a functional luciferase messenger RNA may be or may not be transcribed (**Fig. IV-2c**). The T-DNA integrates into a repeated genomic DNA region in which the promoterless luciferase gene is inactivated and, therefore, no luciferase messenger RNA is synthesized and luciferase activity can not be detected in transgenic plants (**Fig. IV-2d**). Solid line with arrows indicate the direction of gene transcription; Dotted line with arrows indicate the transcription may or may not take place; G, a plant gene; P, a plant promoter; RS, a repeated sequence in a plant genome. *LUX*, a promoterless *lux* gene linked to the right border of T-DNA; *HPT*, hygromycin phospho-transferase for selection of transgenic plants.


(a)
P
G
RS
(b)
P
LUX
HPT
G
RS
Transferred DNA
(c)
P
G
LUX
HPT
G
RS
Transferred DNA
(d)
P
G
RS
LUX
HPT
RS
Transferred DNA

The material involved in page 118 has been removed because of copyright restrictions. This page contains a legend of photographs which has been published. Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Plant Mol. Biol. Rep. 10, 358.

The material involved in page 119 has been removed because of copyright restrictions. This page contains photographs which have been published. Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Plant Mol. Biol. Rep. 10, 358.

The material involved in page 120 has been removed because of copyright restrictions. This page contains a legend of a table which shows the distribution of luciferase gene expression in transgenic tobacco plants. Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Plant Mol. Biol. Rep. 10, 357.

The material involved in page 121 has been removed because of copyright restrictions. This page contains a table which shows the distribution of luciferase gene expression in transgenic tobacco plants. Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Plant Mol. Biol. Rep. 10, 357.

Chart IV-1. Differential expression of the promoterless bacterial luciferase marker gene in transgenic tobacco plants. Two homozygous independent transgenic tobacco plants containing the plant insertional vector pPCVG *luxA&B* were obtained with *Agrobacterium*-mediated transformation and both of them showed constitutively luciferase activity in transgenic plants. Plant tissues were taken from the top, middle, bottom leaves and stems, and from roots of transgenic plants at the similar developmental stage (before flowering). Plant tissue (50 mg of fresh weight) was homogenized in 500 µl of luciferase reaction buffer (lux buffer) and assayed for the luciferase activity in a luminometer. The luciferase activity in this chart was calculated based on the average LU (Light Unit) detected in homogenates of three samples for each individual plant and three individual plants for each transgenic plant. Luciferase activities are expressed as LU mg $^{-1}$ protein where 1 LU is equal to 1.2×10^6 protons sec^{-1}.

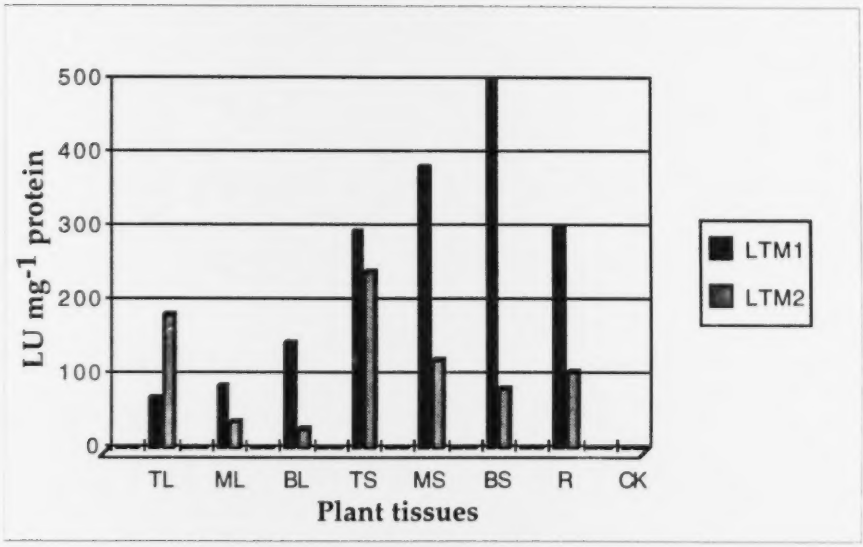
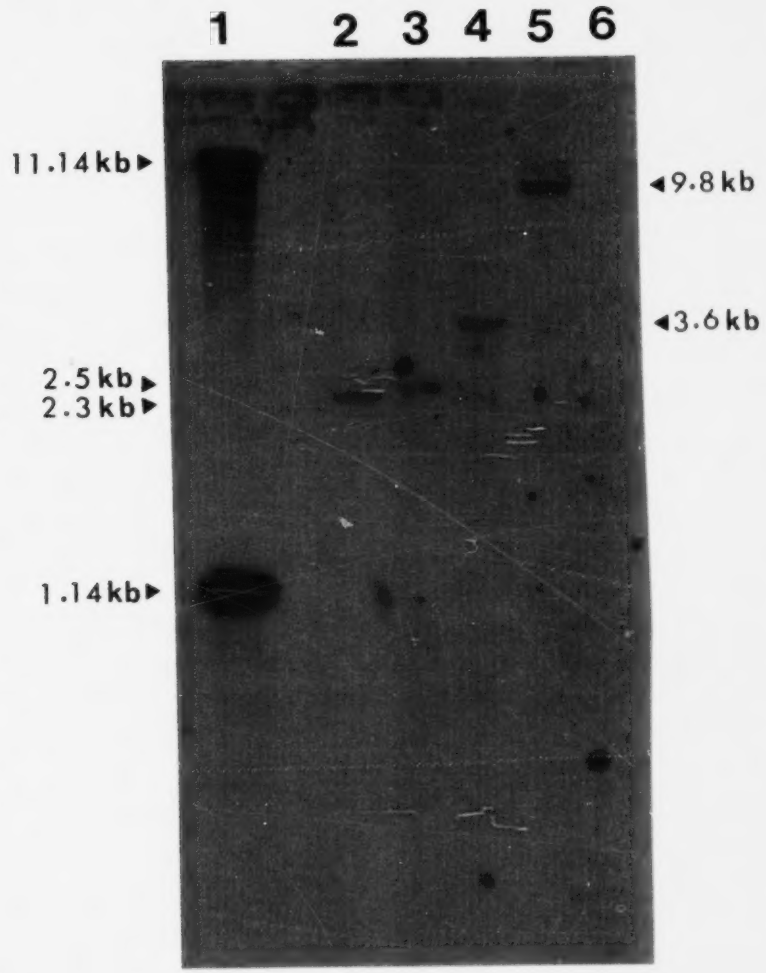

500
400
300
200
100
0
LU mg-1 protein
TL
ML
BL
TS
MS
BS
R
CK
Plant tissues
LTM1
LTM2

Fig. IV-4. Southern blot analysis of plant 12-1 containing plant insertional vector pPCVG *luxA&B* with ^{32}P-radio-labeled *lux A* DNA fragment as a probe. Genomic DNA from plant 12-1 (10 µg each) was digested with Xba I / Hind III (**lane 2**), Xba I (**lane 3**), EcoR I (**lane 4**) and Hind III (**lane 5**).and the DNA fragments were seperated by electrophoresis in TBE buffer. After denaturation, the DNA fragments were transferred to a nylon membrane (see materials and methods). The membrane was hybridized with a radio-labeled *lux A* as a probe. Positive control, pPCVG *lux A&B* was digested with BamH I (**lane 1**) and negative control, plant SR1 genomic DNA (10 µg) was digested with Hind III only (**lane 6**). Solid arrows indicate the size of hybridization bands, and kb represents kilobase.


1
2
3
4
5
6
11.14kb
9.8kb
3.6kb
2.5kb
2.3kb
1.14kb

The material involved in page 126 has been removed because of copyright restrictions. This page contains a legend of a diagram which has been published. Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Plant Mol. Biol. Rep. 10, 355.

The material involved in page 127 has been removed because of copyright restrictions. This page contains a diagram which has been published. Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Plant Mol. Biol. Rep. 10, 355.

Fig. IV-6. Direct amplification of the flanking plant DNA fragment linked with the T-DNA tag from genomic DNA isolated from 12-1 plant by inverse Polymerase Chain Reaction. After PCR reaction, 10 µl from each reaction was loaded on a 0.8% agarose gel for electrophoresis and then the gel was stained with EtBr. Lambda / Hind III molecular weight markers (**lane 1 and 6**), plant 12-1 genomic DNA digested with Xba I followed by self-ligation (**lane 2**), plant 12-1 genomic DNA digested with Xba I without ligation (**lane 3**), plant SR1 genomic DNA digested with Xba I followed by self-ligation (**lane 4**) and PCR reaction control pEM provided by the manufacture (**lane 5**). Solid arrows indicate the size of amplified PCR fragments.

1 2 3 4 5 6
1.384 kb
0.65 kb

Chapter V

Molecular Characterization of a Putative Plant Promoter Responsible for Mediating Gene Expression in Tobacco Flowers

1. INTRODUCTION

Flower development in higher plants is a complex process (Grasser, 1991). The molecular and physiological events that control these processes are not well understood (Drews and Goldberg, 1989). Hybridization experiments with tobacco floral organ system mRNA populations (petal, anther and ovary) showed that each organ expresses a large number of organ specific genes (Kamalay and Goldberg, 1980). According to these researchers, both the anther and ovary contain approximately 10,000 diverse mRNAs which are not detectable in heterologous organ system mRNA or nuclear RNA populations (Kamalay and Goldberg, 1980; 1984).

Much progress toward understanding the mechanisms controlling floral development has been made through the analysis of mutants in *Arabidopsis* and *Antirrhinum* (Coen and Meyerowitz, 1991). A number of genes which are affected by these mutations have been cloned, including *AG* in *Arabidopsis* (Yanofsky *et al.*, 1990) and *DEFA* in *Antirrhinum* (Sommer *et al.*, 1990). DNA sequence analysis revealed that both *AG* and *DEFA* genes encode proteins with a conserved 58 amino acid domain called the MADS box (Yanofsky *et al.*, 1990; Sommer *et al.*, 1990). The products of the *AG* and *DEFA* genes are sequence-specific DNA binding proteins (Mueller and Nordheim, 1991; Schwarz-Sommer *et al.*, 1992). These proteins may affect floral development by regulating patterns of gene transcription via interaction with target sequences located at 5'-flanking region of a floral gene.

Taken together, previous studies indicate that the regulation of gene expression in flower development takes place mostly at the level of transcription (Meyerowitz *et al.*, 1989; Schwarz-Sommer *et al.*, 1990; Coen and Meyerowitz, 1991; Grasser, 1991; Schwarz-Sommer *et al.*, 1992). The transcriptional regulation is controlled by groups of *trans*-acting factors and

cis-acting elements, often located within 5'-flanking regions of genes (Benfey and Chua, 1989; Grasser, 1991; Kuhlemeier, 1992). Therefore, isolation and characterization of 5'-flanking regulatory sequences responsible for regulation of gene expression in flower development will provide more information toward understanding these complex processes.

This chapter mainly focus on the molecular analysis of a putative plant promoter element responsible for mediating regulation of gene expression in tobacco flowers. Sequence analysis of this DNA fragment revealed the location of a TATA-like box and three CAAT boxes in the region closely linked to the 5' end of the promoterless *lux A* gene. In addition, a consensus sequence similar to the control region found in other flower specific genes was identified. Further, bioluminescence was detected in flowers of the transgenic tobacco plants, when this putative promoter region (1.065 bp), fused with the *lux F* marker gene, was reintroduced by transformation into tobacco cells and transgenic plants was regenerated.

MATERIALS AND METHODS

Bacterial strains and DNA manipulation

Competent cells were prepared with *E. coli* strains *DH5α* and *S17-1* as described previously (Inoue *et al.*, 1990). A plant promoter test vector, pTV-1p, was mobilized from *S17-1* into *A. tumefaciens* strain *GV3101 pMP 90RK* according to the method described previously (Koncz and Schell, 1986). The correct transconjugant was selected according to the method described in Chapter III of this thesis and used to transform tobacco plants. DNA manipulation involved in the vector construction was carried out by using standard protocols (Sambrook *et al.*, 1989).

Plant transformation and regeneration

Nicotiana tabacum cv. SR1 leaf-discs, 8-mm diameter, were harvested at the 3-4 leaf stage, and were grown on MS medium in sterile jars. After soaking for 15 minutes with log phase growth *A. tumefaciens* strain *GV 3101 pMP 90RK* containing a plant test vector, pTV-1p, the leaf discs were incubated on MS medium at 25°C, without antibiotic selection for two days in the dark. The leaf discs were transferred to MS shooter medium containing 20 mg/l of hygromycin for selection of the plant insertional vector and 100 mg/l of kanamycin for selection of the promoter test vector. Kanamycin-resistant shoots, which appeared after three to four weeks, were excised from the leaf discs and rooted in jars containing MS medium without hormones but containing kanamycin selection. Two to three weeks later, the transformed plantlets with substantial roots were grown to maturity in soil in the greenhouse under a 12-hour photoperiod (Horsch *et al.*, 1985).

Screening transformed plants

Luciferase activity of homogenate from transformed tobacco plants was carried out essentially as described previously (Olsson et al., 1988; Langridge *et al.*, 1991). Plant tissues (50 mg) were harvested form transformed plants greenhouse grown under a 12-hour photoperiod. The samples were ground in 500 µl of lux-buffer (50 mM Na-phosphate, pH 7, 50 mM β-mercaptoethanol, 0.4 M sucrose) (Koncz *et al.*, 1987), homogenate was distributed into two microcentrifuge tubes, one, 50 µl, for protein assay (Bradford, 1976) and the other, 450 µl, for analysis of luciferase activity. The sample for luciferase assay was placed in a luminometer (Turner TD-20e), and 20 µl decanal substrate (Sigma), as a sonicated 1:1000 dilution in ddH_2O, and 500 µl light-reduced FMN in tricine buffer (200 mM tricine pH 7) were

injected into the luminometer to start the reaction. The high peak of the light produced during the first 10 sec of the reaction was taken as the luciferase activity, where 1 light unit (LU) is equal to 1.2×10^6 photons sec^{-1} (Koncz *et al.*, 1987).

Subcloning and DNA sequencing

The amplified iPCR fragments were purified on an 0.8% agarose gel, eluted from the gel and subcloned into pBluescript (Stratagene) vectors (SK). Subsequently, inserts were analyzed by restriction endonuclease enzyme mapping, Southern hybridization and sequence analysis. Restriction enzyme digestion and plasmid manipulations were according to standard procedures (Sambrook *et al.*, 1989). Double-stranded DNA sequencing was performed using both cycle sequencing Dye-primer and Dye-terminator Kits (Applied Biosystems, Inc.) on an ABI DNA Sequencer following the protocol provided by the manufacturer (Applied Biosystems, California). Sets of undirectional deletions were generated by digestion with exonuclease III as described previously (Henikoff *et al.*, 1984). Both strands were sequenced for each subclone.

RESULTS AND DISCUSSION

Subclones and sequencing strategies

T-DNA tagging experiments with a specialized plant insertional vector, pPCVG *lux A&B*, containing a promoterless bacterial luciferase gene yielded various insertional mutant plants (Jiang *et al.*, 1992; 1993b). One insertional mutant, plant 12-1, showed the luciferase gene predominantly expressed in flowers. The collected photon image of a flower from plant 12-1 showed the high level of luciferase activity in the corolla (Fig. V-1). Using iPCR methods

(Ochman *et al.*, 1988), a 5'-flanking DNA sequence was directly isolated from the plant 12-1 genome (see Chapter IV). This PCR product was subcloned in the Sma I site of pBluescript (SK) (Stratagene). Independent recombinant clones were randomly recovered from white colonies grown on LB/Agar medium containing 100 mg of ampicillin and resulting colonies were subjected to miniprep analysis. Fig. V-2 shows the results of restriction enzyme analysis of two independent clones, pBS-1 and pBS-2, containing the same size inserts that were released by Xba I / Sal I, and Hind III / Sal I. For pBS-1, six serial deletions in the direction from 5' to 3' were obtained by Exo III-mediated deletion experiments (Fig. V-3a; 3b). The sequencing strategies of the FEP (the putative flower expression promoter) are shown in Fig. V-4. Double-stranded DNA sequencing was performed with -21M13 forward and reverse primers (ABI). The complementary strand of pBS-1 and pBS-2 were sequenced by using site-specific primers. The sequence data confirmed that both clones, pBS-1 and pBS-2, contained the same insert of 1384 bp, including the flanking plant DNA and the primer sequences (Fig. V-5a). The complete sequence of the fragment (1073 bp) is listed in Fig. V-5b, including the flanking plant DNA (1030 bp), the T-DNA right border (11 bp) and the start codon of the *lux A* gene.

Sequence analysis of a putative plant promoter

DNA sequence analysis of this 1.384 kb fragment revealed that the two PCR primers were located on the ends of the fragment. The T-DNA right border sequence, GTTTACCCTCA (-40 to -30) was found in a position 29 bp upstream of the ATG start codon of the *lux A* gene. This result indicated that there was a 2 bp deletion (AA) between the T-DNA right border and the ATG start codon of the *lux A* gene. Compared with the T-DNA right border

sequence, a 14 nucleotide deletion was identified within the T-DNA right border after integrated in the plant genome (Fig. V-6).

The 14 nucleotide deletion found in the right border sequence is different from previous reports, one of which suggested that the right border of T-DNA was much better preserved than the left one (Hohn *et al.*, 1991). Since the left border data in this experiment are currently not available, it is impossible to compare them. However, compared with published data of thirteen independent clones from both *Arabidopsis* and tobacco transgenic plants, a fourteen nucleotide deletion at the right border of T-DNA is the largest reported (Mayerhofer *et al.*, 1991; Gheysen *et al.*, 1991).

A Hind III site was found in a position 173 bp downstream of a Xba I site. This is consistent with the result from Southern hybridization (see Fig. IV-4). The sequence was highly AT-rich (60.29%) and a palindromic sequence, TTAAAAAAAT-AC-ATTTTTTTAA (-657 to -636), was observed. This palindromic sequence contained two 10 bp inverted repeats separated by 2 bp. The function of this palindromic structure is not clear, but previous studies of the *lac* operator in *E. coli* have suggested that the symmetrical structure may block transcription (repressor) (Jacob and Monod, 1961; Gilbert and Muller-Hill, 1966).

Consensus sequences

Mutational analyses of eucaryotic promoters have defined two classes of DNA sequences that are important for transcriptional regulation of gene expression (Kuhlemeier, 1992). First there is the TATA box, or functionally related sequences that bind the RNA polymerase complex and determine the transcription start site (Carcamo *et al.*, 1990). The second class of DNA elements includes many upstream sequence elements (USEs) which function at variable distances from the TATA box. These include the CAAT box, GC

motif and many consensus sequences. These sequences enhance or repress transcription under specified cellular or environmental conditions (Mitchell and Tijan, 1989).

The sequence data revealed the location of a putative TATA-like sequence, TATAACA, (-39 to -33) and three putative CAAT boxes (-94 to -91, -64 to -61, and -57 to -54) in the region closely linked to the 5' end of the promoterless *lux A* gene. For most pol II transcribed genes, the TATA box is well known to interact with the basic transcription initiation factor TFIID which functions in organizing the assembly of the transcription complex, including factors TFIIA, TFIIB, TFIIE, TFIIF and pol II onto the promoter (Greenblatt, 1991). Three putative CAAT boxes were found in the region upstream of the TATA box. The upstream sequence elements (USEs) such as CAAT boxes can control both induction and repression of transcriptional activity depended on the different cell types and the trans-acting factor that it is associated with (Lewin, 1990).

In addition, a search for consensus binding sites of known transcription factors revealed a similar sequence of SRF box (a target sequence of a mammalian transcription factor), CCATAATTGA (-269 to -260) which represents the consensus of the binding sites for floral MADS-box-containing transcription factors (Meyerowitz *et al.*, 1989; Schwarz-Sommer *et al.*, 1990; 1992; Coen and Meyerowitz, 1991). The putative promoter containing a DNA binding site of transcription factors of floral homeotic genes indicated that this promoter may control a target gene which is involved in flower development. Genebank searches for homologues of this putative plant promoter demonstrated no sequence homologies above the 50% level. Further investigation of this promoter, such as mutation or deletion analysis

and isolation of its corresponding gene is necessary to provide a clear picture about its function in flower development.

Construction of a promoter test vector, pTV-1p

In order to confirm that the flanking plant DNA fragment was responsible for regulation of the *lux* marker gene expression in tobacco flowers, a promoter test vector, pTV-1p, was constructed. Figure V-7 shows the construction procedure of the promoter test vector, pTV-1p. A 1.065 kb insert was released from pBS-1 after digestion with Xba I and Sal I. The Xba I / Sal I fragment was cloned into a plant expression vector pPCV701 *luxFM2* (Jiang *et al.*, 1993a) by replacing the mas promoter (Koncz *et al.*, 1987). This formed the promoter test vector, pTV-1p, which contained the putative promoter, the coding sequence of the *lux F* marker gene (Olsson *et al.*, 1989; Escher *et al.*, 1989), the termination signal from the octopine synthase polyadenylation sequence (De Greve *et al.*, 1982) and kanamycin resistant gene (*npt II*) (Bevan *et al.*, 1983). The promoter-*lux F* gene fusion and the termination signal are supposed to form the transcriptional apparatus when integrated into the plant genome (Kuhlemeier, 1992). The use of the luciferase marker gene provides an indication of T-DNA integration, and offers a convenient tool to monitor promoter activity in transgenic plants. The *npt II* gene provides a convenient selectable marker for selection and identification of transformed plants. There was no strong enhancer sequences adjacent to the transcriptional unit in the vector that could influence gene expression patterns (Koncz *et al.*, 1989).

The Putative promoter mediated luciferase gene expression in transgenic Plants

The promoter test vector, pTV-1p was mobilized into *A. tumefaciens* strain *GV 3101 pMP 90RK* and used to transform tobacco leaf discs. Twenty-three kanamycin-resistance transgenic tobacco plants were regenerated and root, leaf, stem and flower tissues from these transformed plants were screened for luciferase activity *in vitro* and *in vivo* at different stages during plant development. The *lux* gene expression was detected only in flowers in nine of these plants.

The level of *lux* gene expression was about twenty times lower than that in plant 12-1 (Chart V-1). The different level of the *lux* marker gene expression in the 12-1 plant and retransformed plants reflects complicated processes of regulation of gene expression. This result suggests that the length of the promoter and position effects may play a role (Benfey and Chua, 1989).

The isolation, molecular characterization and retransformation experiments confirm that the isolated putative promoter does direct flower specific gene expression. The 5'-flanking plant DNA linked to the *lux A* gene is responsible for mediating gene expression in tobacco flowers. The sequence data reveals that this putative promoter element contains several consensus sequences common to other plant promoters. The SRF-like box, CAAT boxes and a palindromic structure may play roles for the flower-speficic gene expression. These results suggest that this putative promoter element may control a gene involved in flower development in tobacco. Further investigations such as point mutation, deletion analysis and gel retardation assays are necessary for elucidation of function and involvement of this promoter in flower development. Using this promoter element as a

hybridization probe to isolate its corresponding gene from a tobacco genomic library will provide more information about the role of this promoter and its gene product during flower development. Plant promoters have many potential applications in both basic and applied research. For example, they can be used to study regulation of gene expression, to identify a nearby gene by chromosome walking, as molecular markers in RFLP experiments and to express a foreign gene of interest in higher plants (Chang *et al.*, 1988; Mariani *et al.*, 1990).

Additional questions relating to chapter V

Possible reasons for the decreased expression level observed with the cloned promoter

The retransformed tobacco plants with the cloned promoter showed the similar gene expression pattern compared to that in the primary transgenic plant. However, the level of *lux* expression in retransformed tobacco plants with the cloned promoter was dramatically decreased in comparison to that in the primary transgenic plant. Most genes whose expression has been studied in transgenic plants are expressed in generally appropriate patterns with respect to cell and organ specificity, developmental timing, and to environmental cues (Benfey and Chua, 1989; Kuhlemeier, 1992). However, the level of expression can vary over an extremely wide range (Dean *et al.*, 1988; Hobbs *et al.*, 1990 and 1993). As was discussed in chapter II, the regulation of gene expression is a very complex process. Low level of the *lux* gene expression with a cloned promoter may be caused by several factors, including length of the cloned promoter, different transcriptional fusions between the insertional vector and the promoter test vector, translational regulation of bacterial luciferases.

Firstly, because the cloned promoter is only 1065 bp in length, it is possible that enhancer-like elements which reside in the native promoter were deleted. In the primary transformant, the T-DNA integrated downstream of the plant promoter resulting in a transcriptional fusion between the promoterless luciferase gene and the native promoter. The activation of the *lux* gene then was detected *in vivo* by low-light video image analysis. Therefore, the *lux* gene expression pattern and level was the reflection that of the native gene at least at the transcriptional level. If this is the case, the decreased expression level observed with the cloned promoter could be explained.

Secondly, there is a 35S promoter of the cauliflower mosaic virus located at 3' end of the *lux A* gene in the insertional vector used in this experiment (see Fig. IV-1). There is no such strong enhancer in the promoter test vector (see Fig. V-6). The 35S promoter of the cauliflower mosaic virus contains a complex enhancer element that has the ability to influence gene expression over a long distance in either orientation when present upstream or downstream of genes (Kay *et al.*, 1987; Nagy *et al.*, 1987; Fang *et al.*, 1989). A similar observation was also reported in a T-DNA tagging experiment with the 35S promoter of the cauliflower mosaic virus (Koncz *et al.*, 1989). Therefore, another possible explanation for different levels of the *lux* gene expression observed between the retransformed plant and the primary transformant may be the presence of the enhancer of the 35S promoter in the insertional vector.

Finally, translational regulation of the reporter gene may play a role in the decreased expression level observed in the retransformed plant compared to the primary transformant. In the insertional vector, the *lux A* and *lux B* genes were used separately (see Fig. IV-1), whereas in the promoter test

vector, the *lux F* (*lux A* and *lux B* fusion) gene was used (see Fig. V-6). It has been reported that the the luciferase $\alpha\beta$ fusion protein had only 80% of the wild-type activity (Olsson *et al.*, 1989).

The above discussion only focuses on the transcriptional and translational regulation and its possible relationship to the decreased expression level observed with the cloned promoter. In fact, as was discussed in chapter II, section 5., even with transgenic plants generated from an identical T-DNA construct, the level of transgene expression can vary a hundred fold (136-175-fold) (Peach and Velten, 1991). Based on their observation on transgene expression, Peach and Velten (1991) suggested that a majority of detectable transformation events results in very low expression. This observation indicates that factors other than those discussed above also influence the expression of the transgene.

What additional evidence would be needed to prove that the cloned DNA is indeed a promoter?

Insertional activation of a reporter gene for studying the regulation of gene expression followed by isolation of its flanking sequences has been successfully used both in animals (O'Kane and Gehring, 1987; Allen *et al.*, 1988; Bellen *et al* 1989) and plants (Ander *et al.*, 1986; Terri *et al.*, 1986; Koncz *et al.*, 1989; Herman *et al.*, 1990; Kertbundit *et al.*, 1991; Topping *et al.*, 1991; Lindsey *et al.*, 1993). These previous studies demonstrated that expression patterns of insertional activation reflect the regulatory properties of flanking sequences. Using this strategy, promoter- or enhancer-elements have been isolated from both animals and plants (for review see Walden *et al.*, 1991; Lijisebettens *et al.*, 1991; Koncz *et al.*, 1992). A similar strategy using a promoterless bacterial luciferase gene, led to isolation of a flanking plant DNA which is responsible for activation of the promoterless *lux* gene

expressed in tobacco flowers. Sequence analysis of the cloned flanking DNA followed by functional test in the retransformed plants demonstrated that the cloned DNA has several consensus sequences, such as TATA-like box and CAAT boxes, and direct a reporter gene expression in an organ-specific fashion. However, in order to prove that the cloned DNA is indeed a promoter, additional evidence is needed.

A promoter is the region immediately up-stream of each transcription initiation site that is closely linked to the down-stream coding region through a short untranslated region (Lewin, 1990). An typical processed RNA transcript in eukaryotes consists of a sequence extending from the transcription start site at the 5' to the terminator at the 3' ends (minus introns) (Kuhlemeier, 1992). Therefore, mapping the 5' end of the untranslated region in a transcript made from the cloned DNA-*lux* gene fusion will not only reveal the transcription start site in DNA, it will provide evidence whether there is a direct linkage between the cloned DNA and the *lux* gene.

The transcription start site in the mRNA molecule made from the cloned DNA-*lux* gene fusion can be determined by using the primer extension (Mcknight *et al.*, 1981) or the endonuclease S1 mapping techniques (Berk and Sharp, 1977). In the primer extension technique, an oligonucleotide (primer) which is complementary to the *lux* mRNA is prepared and end-labeled. After it is hybridized to the mRNA, the primer can be extended by the enzyme reverse transcriptase until it reaches the first nucleotide of the mRNA. The length of this extension product can then be determined accurately by gel electrophoresis by comparing with a sequencing ladder derived from the DNA with the same oligonucleotide primer. The location of the TATA box also can be determined. The use of S1 to map the

start site of the *lux* mRNA begins with the preparation of an end-labeled, short (approximately 100 nucleotides), single stranded DNA, that encodes the general region of the mRNA start site. This end-labeled, single stranded DNA is used to hybridize with the mRNA, and any unpaired nucleic acid is then digested with S1 endonuclease. Denaturation leaves a labeled DNA fragment whose length accurately marks the distance of the starting nucleotide of the mRNA.

Using the cloned DNA as a hybridization probe to pull out its down-stream region from a genomic library will reveal whether the cloned DNA is indeed linked to a coding region of a plant gene. However, primer extension experiments still needed to determine where the mRNA initiated.

BIBLIOGRAPHY

Allen, N.D., Cran, D.G., Barton, S.C., Hettle, S., Reik, W. and Surani, M.A. (1988) Transgenes as probes for active chromosomal domains in mouse development. *Nature* 333, 852-55.

Andre, D., Schell, J., Van Montagu, M. and Hernalsteens, J.P. (1986) Gene tagging in plants by a T-DNA insertion that generates APH(3')II plant gene fusions. *Mol. Gen. Genet.* 204, 512-18.

Bellen, H.I., O'Kane, C.J., Wilson, C., Grossniklaus, U., Pearson, R.K. and Gehring, W.J. (1989) P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. *Genes & Dev.* 3, 1288-1300.

Benfey, P.N. and Chua, N.H. (1989) Regulated genes in transgenic plants. *Science* 242, 174-81.

Berk, A.J. and Sharp, P.A. (1977) Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. *Cell* 12, 721-32.

Bevan, M., Flavell, P.B. and Chilton, M.D. (1983) A chimeric antibiotic resistance gene as a selectable marker for plant cell transformation. *Nature* 304, 184-187.

Bradford, M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein, utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248-51.

Carcamo, J., Maldonado, E., Cortes, P., Ahn, M.H., Ha, I., Kasai, Y., Flint, J. and Reinberg, D. (1990) A TATA-like sequence located downstream of the transcription initiation site is required for expression of an RNA polymerase II transcribed gene. *Genes & Dev.* 4, 1611-22.

Chang, C., Bowman, J.L., Dejohn, A.W., Lander, E.S. and Meyerowitz, E.M. (1988) Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 85, 68-6860.

Coen, E.S. and Meyerowitz, E.M. (1991) The war of the whorls: Genetic interactions controlling flower development. *Nature* 353, 31-37.

De Greve, H., Dhaese, P., Seurinck, J., Lemmers, M., Van Montagu, M. and Schell, J. (1982) Nucleotide sequence and transcript map of the *Agrobacterium tumefaciens* Ti plasmid-encoded octopine synthase gene. *J. Mol. Appl. Gen.* 1, 499-511.

Dean, C., Jones, J., Favreau, M., Dunsmuir, P. and Bedbrook, J. (1988) Influence of flanking sequences on variability in expression levels of an introduced gene in transgenic tobacco plants. *Nucl. Acids Res.* 16, 9267-83.

Drews, G.N. and Goldberg, R.B. (1989) Genetic control of flower development. *TIG.* 5, 256-61.

Escher, A., O'Kane, D.J., Lee., J. and Szalay, A.A. (1989) Bacterial luciferase αβ fusion protein is fully active as a monomer and highly sensitive *in vivo* to elevated temperature. *Proc. Natl. Acad. Sci. USA* 86, 6528-32.

Fang, R.-X., Nagy, F., Sivasubramaniam, S. and Chua, N.-H. (1989) Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell* 1, 141-50.

Gasser, C.S. and Fraley, R.T. (1989) Genetically engineering plants for crop improvement. *Science* 244, 1293-99.

Gilbert, W. and Muller-Hill, B. (1966) Isolation of the Lac repressor. *Proc. Natl. Acad. Sci. USA* 56, 1891-98.

Grasser, C.S. (1991) Molecular studies on the differentiation of floral organs. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42, 621-49.

Gheysen, G., Villarroel, R. and Van Montagu, M. (1991) Illegitimate recombination in plants: a model for T-DNA integration. *Genes Dev.* 5, 287-97.

Greenblatt, J. (1991) Roles of TFIID in transcriptional initiation by RNA polymerase II. *Cell* 66, 1067-70.

Henikoff, S. (1984) Bidirectional digestion with exonuclease III creates breakpoints for DNA sequencing. *Gene* 28, 351-59.

Herman, L., Jacobs, A., Van Montagu, M. and Depicker, A. (1990) Plant chromosome/marker gene fusion assay for study of normal and truncated T-DNA integration events. *Mol. Gen. Genet.* 224, 248-56.

Hobbs, S.L.A., Kpodar, P. and DeLong, C.M.O. (1990) The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. *Plant Mol. Biol.* 15, 851-64.

Hobbs, S.L.A., Warkentin, T. and DeLong, C.M.O. (1993) Transgene copy number can be positively or negatively associated with transgene expression. *Plant Mol. Biol.* 21, 17-26.

Hohn, B., Koukolikova-Nicola, Z., Durrenberger, F., Bakkeren, G. and Koncz, C. (1991) The T-DNA on its way from *Agrobacterium tumefaciens* to the plants. *In Advances in Molecular Genetics of Plant-Microbe Interactions* 1, 19-27.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. (1985) A simple and general method for transferring genes into plants. *Science* 227, 1229-31.

Jacob, F. and Monod, J. (1966) Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318-56.

Jiang, C., Escher, A. and Szalay, A.A. (1993a). Alteration of a luciferase based binary plant transformation vector for simultaneous expression of desired

gene products and bacterial luciferase in transgenic plants. Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds.) John Wiley & Sons. 127-31.

Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Identification of plant genes *in vivo* by tagging with T-DNA border-linked luciferase genes followed by inverse polymerase chain reaction amplification. *Plant Mol. Biol. Rep.* 10, 345-61.

Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1993b) Isolation of developmentally regulated promoter elements from tobacco by T-DNA insertional activation of promoterless luciferase marker gene. Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds.) John Wiley & Sons. 217-21.

Kamalay, J.C. and Goldberg, R.B. (1980) Regulation of structural gene expression in tobacco. *Cell* 19, 935-46.

Kamalay, J.C. and Goldberg, R.B. (1984) Organ-specific nuclear RNAs in tobacco. *Proc. Natl. Acad. Sci. USA* 81, 2081-85.

Kay, R., Chan, A., Daly, M. and Mcpherson, J. (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. 236, 1299-1302.

Kertbundit, S., De Greve, H., DeBoeck, F., Van Montagu, M. and Hernalsteens, J.P. (1991) *In vivo* random β-glucuronidase gene fusion in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 88, 5212-16.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Korber, H., Redei, G.P. and Schell, J. (1989) High-frequency T-DNA-mediated gene tagging in plants. *Proc. Natl. Acad. Sci. USA* 86, 8467-71.

Koncz, C., Nemeth, K., Redei, G. and Schell, J. (1992) T-DNA insertional mutagenesis in *Arabidopsis*. *Plant Mol. Biol.* 20, 963-976.

Koncz, C., Olsson, O., Langridge, W.H.R., Schell, J. and Szalay, A.A. (1987) Expression and assembly of functional bacterial luciferase in plants. *Proc. Natl. Acad. Sci. USA* 84, 131-35.

Koncz, C. and Schell, J. (1986) The promoter of the T1-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* 204, 383-96.

Kuhlemeier, C. (1992) transcriptional and post-transcriptional regulation of gene expression in plants. *Plant Mol. Biol.* 19, 1-14.

Langridge, W.H.R., Escher, A.P. and Szalay, A.A. (1991) Measurement of bacterial luciferase as a reporter enzyme *in vivo* in transformed bacteria, yeast and plant cells and in transgenic plants. *Technique* 3, 99-108.

Lewin, B. (1990) Genes IV. *Oxford University Press and Cell Press.* 543-77.

Lijsebettens, M.V., den Boer, B., Hernalsteens, J.-P. and Van Montagu, M. (1991) Insertional mutagenesis in *Arabidopsis thaliana*. *Plant Sci.* 80, 27-37.

Lindsey, K., Wei, W., Clarke, M.C., McArdle, H.F., Rooke, L.M. and Topping, J.F. (1993) Tagging genomic sequences that direct transgene expression by activation of a promoter trap in plants. *Transgenic Res.* 2, 33-47.

Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Bakkeren, G., Crameri, A., Angelis, K., Redei, G.P., Schell, J., Hohn, B. and Koncz, C. (1991) T-DNA integration: a model of illegitimate recombination in plants. *EMBO J.* 10, 697-704.

Mariani, C., De Beuckeleer, M., Truettner, J., Leemans, J. and Goldberg, R.B. (1990) Induction of male sterility in plants by a chimeric ribonuclease gene. *Nature* 347, 737-41.

McKnight, S.L., Gavis, E.R. and kingsbury, R. (1981) Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upsteam control region. *Cell* 25, 385-98.

Meyerowitz, E.M., Smith, D.R. and Bowman, J.L. (1989) Abnormal flowers and pattern formation in floral development. *Development* 106, 209-17.

Mitchell, P.J. and Tjian, R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245, 371-78.

Nagy, F., Boutry, M., Hsu, M.-Y., Wong, M. and Chua, N.-H. (1987) The 5'-proximal region of wheat *Cab-1* gene contains a 268-bp enhancer-like sequence for phytochrome response. *EMBO J.* 6, 2537-42.

Ochman, H., Gerber, A.S. and Hartl, L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* 120, 621-23.

Olsson, O., Escher, A.P., Sandberg, G., Schell, J., Koncz, C. and Szalay, A.A. (1989) Engineering of monomeric bacterial luciferases by fusion of *luxA* and *luxB* genes in *Vibrio harveyi*. *Gene* 81, 335-47.

Olsson, O., Koncz, C. and Szalay, A.A. (1988) The use of the *lux A* gene of the bacterial luciferase operon as a reporter gene. *Mol. Gen. Genet.* 215, 1-9.

Peach, C. and Velten, J. (1991) Transgene expression variability (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. *Plant Mol. Biol.* 17, 19-60.

Sambrook, J., Maniatis, T. and Fritsch, E.F. (1989) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sanger, F., Niklen, S. and Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463-67.

Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. and Sommer, H. (1990) Genetic control of flower development by homeotic genes in *Antirrhinum majus*. *Science* 250, 931-36.

Schwarz-Sommer, Z., Hue, I., Huijser, P., Flor, P.J., Hansen, R., Tetens, F., Lonnig, W.E., Saedler, H. and Sommer, H. (1992) Characterization of the

Antirrhinum floral homeotic MADS-box gene *DEFICIENS*: Evidence for DNA binding and autoregulation of its persistent expression throughout flower development. *EMBO J.* 11, 251-63.

Sommer, H., Beltran, J.P., Huijser, P., Pape, H., Lonnig, W.E., saedler, H. and Schwarz-Sommer, Z. (1990) *deficiens*, a homeotic gene involved in control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. *EMBO J.* 9, 877-92.

Teeri, TH., Herrera-Estrella, L., Depicker, A., van Montagu, M and Palva, ET. (1986) Identification of plant promoters in situ by T-DNA-mediated transcriptional fusions to the *npt II* gene. *EMBL J.* 5, 1755-60.

Topping, J.F., Wei, W. and Lindsey, K. (1991) Functional tagging of regulatory elements in the plant genome. *Development* 112, 1009-19.

Walden, R., Hayashi, H. and Schell, J. (1991) T-DNA as a gene tag. *Plant J.* 1, 281-88.

Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A. and Meyerowitz, E.M. (1990) The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature* 346, 35-39.

FIGURES AND LEGENDS

Fig. V-1 Bacterial luciferase expression visualized in flower of insertional mutant plant 12-1 detected by low-light image analysis. Two tobacco flowers were placed adjacent to each other in a 9-cm diameter culture dish, decanal substrate was added to the lid and the plate containing the flowers was subjected to low light image analysis for 30 minutes. The flower on the left (**A**) was excised from an nontransformed *N. tabacum* SR1 plant and the flower on the right (**B**) was taken from plant 12-1 which was obtained from a T-DNA tagging experiment using the plant insertional vector pPCVG *luxA&B*. The collected photon image of the flowers showed high levels of luciferase activity in the corolla of the flower from the transgenic plant.

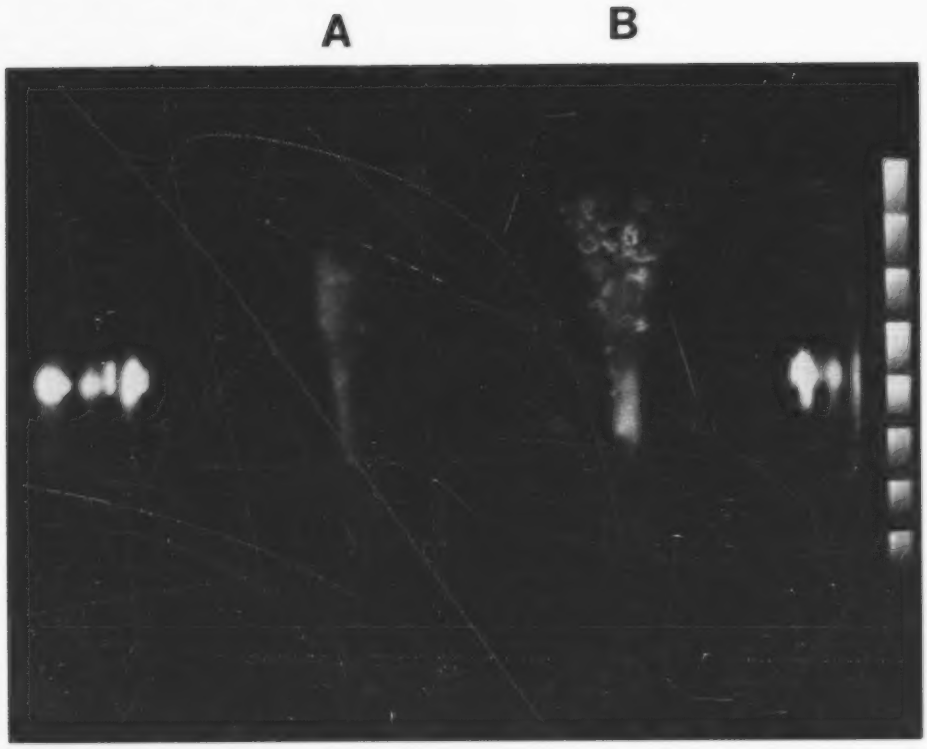

A
B

Fig. V-2. Restriction enzyme analysis of two independent clones, pBS-1 and pBS-2 , containing an insert from inverse PCR. Miniprep plasmid DNA isolated from two independent clones, pBS-1 (**lane 2** and **3**) and pBS-2 (**lane 4** and **5**) were digested to completion and were loaded on a 0.8% agarose mini-gel for electrophoresis. Upon completion, the gel was stained with EtBr. The pBS-1 (**lane 2** and **3**) and pBS-2 (**lane 4** and **5**) were digested with Sal I / Xba I (**lane 2** and **4**) and Sal I / Hind III (**lane 3** and **5**). Lane 1 is the Lambda / Hind III molecular weight marker. The size of the inserts is indicated.

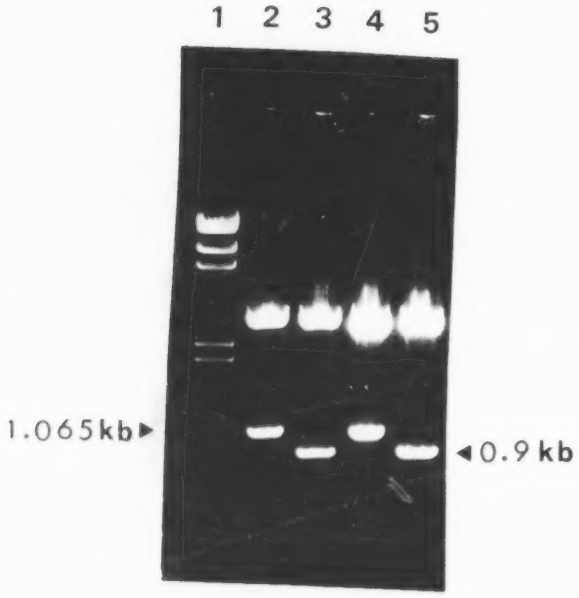
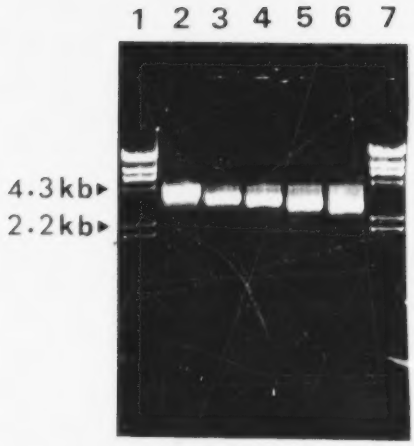

1 2 3 4 5
1.065kb
0.9 kb

Fig. V-3. Generation of unidirectional deletions with exonuclease III. (**a**) Plasmid (pBS-1) DNA (10 μg) was digested to completion with Xba I and Kpn I, respectively, and incubated with Exo III (Sigma) according to the manufacturer's protocol. The DNA samples were taken at 30 second intervals and were loaded on a 0.8% agarose mini-gel for electrophoresis. Lane 1 and lane 7 are Lambda / Hind III molecular weight markers. Lane 2 to lane 6 represent the samples taken from zero second to two min. (**b**) Six random clones were analyzed by restriction enzyme digestion (Kpn I / Sal I) and various length of inserts were obtained. Lane 1 and lane 8 are Lambda / Hind III molecular weight markers. Lane 2 to lane 7 are clones, recovered from the experiment, containing various length of inserts. The size of some inserts is showed.


1 2 3 4 5 6 7
4.3kb
2.2kb

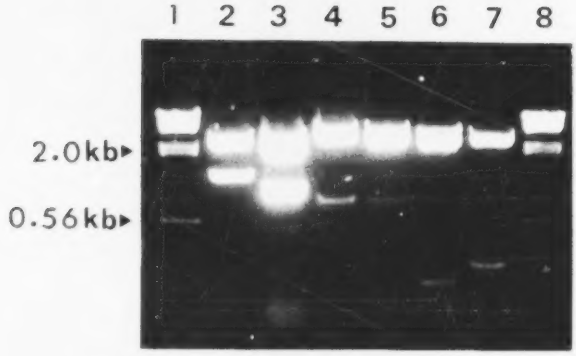

1 2 3 4 5 6 7 8
2.0kb
0.56kb

Fig. V-4. Schematic representation of sequencing strategies of two independent clones, pBS-1 and pBS-2, containing 1.384 kb insert from inverse PCR. For pBS-1, six serial deletions in the direction from 5' to 3' ends were obtained. For the complementary strand of pBS-1 and both strands of pBS-2, site-specific primers were made for carrying out sequencing. Solid arrows indicate the directions of sequencing. The length of the insert is represented by kilobase.

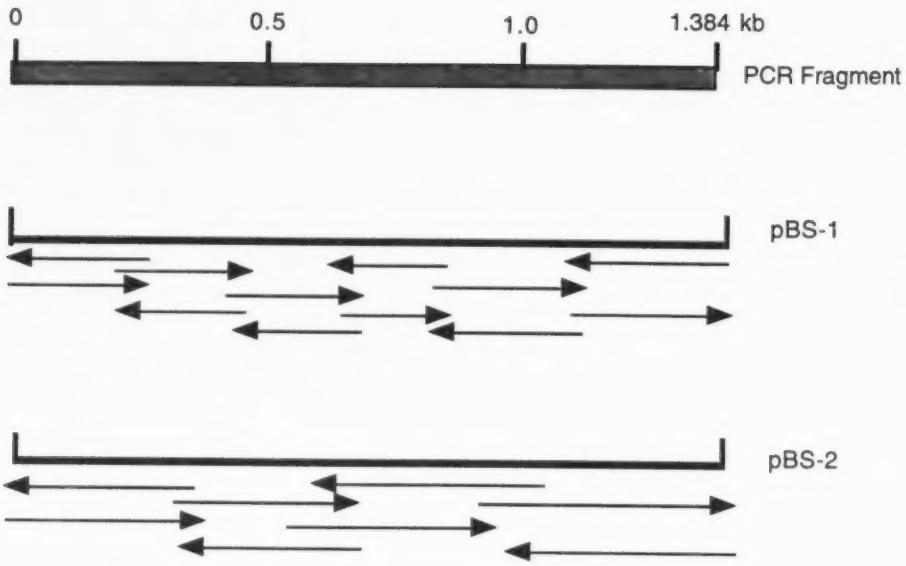
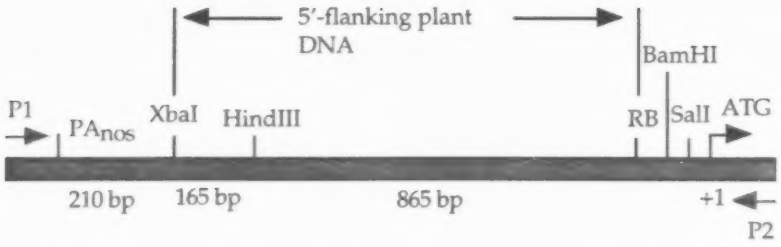

0
0.5
1.0
1.384 kb
PCR Fragment
pBS-1
pBS-2

Fig. V-5. Structure of the PCR fragment and DNA sequence analysis of the FEP element which controls the bacterial luciferase gene predominantly expressed in flowers of plant 12-1. (**a**) The location of two PCR primers (P1 and P2), restriction endonuclease sites and structure of the PCR fragment are marked. Short solid arrows indicate the direction of the inverse PCR. Long solid arrows show the location of the 5'-flanking plant DNA. The length of each fragment is represented by a number of base pair (bp). RB is the T-DNA right border, PAnos is a termination signal from the nonpaline synthase gene of T-DNA and ATG is the translation start codon of the *lux A* gene. (**b**) The ATG translation start codon of the *lux A* gene is shown as underlined and bold character. The ATG start codon of the *lux A* gene is designated as a +1. The right border of the T-DNA, putative TATA-like box and CAAT boxes are bold and italicized. Motifs similar to SRF box are shown as bold, italicized and underlined characters. Palindromic sequence and repeated sequences are underlined. Restriction enzyme sites are marked and underlined.



XbaI

-1070 TCTAGAAAACAAACATCATCTCCAACATAGAGATCATTGCAGTCCACCGT

-1020 TAAAGAAATCCCTAAATCGTATATGTATACATTCAATCTTTTCAGGTTTA

-970 ATGTAAGACTTTGTTTGCACTCTCAACAAGTAAATAGAGAAAAAGAGCAT

HindIII

-920 CTTGGTGCACTAAGCTTCCGCTATGTGCGGGTCTGGGGAAGAGCTGAACA

-870 CAAGGGTCTATCGTACGTAGCCTTACCCTGCATTTTTCAAGAGCTATTTC

-820 ACGGCTGACCAGTGACTCCTGATCACATAGAGCAATTTACTAGTATGTCA

-770 GGCTCCCTCAACAAGTAAATAGACCTCTACATTAATTTTGTGAACAAAGT

-720 TGTTTCAATCTGTTTAAAAAAAATACATTTTTTTAATAACATGTCAGCCAG

-670 TACACAAGGTATCCCGCGTTCAACACAGGGAGAGGCTCGCACTCAAGGAG

-620 AACGATGTACACAGTTTACCCTAATGCAAGTATTAGTGACTGCTTTTGCG

-570 GTTCGAACTCATGAGCTATATGTCATACAGAGACAACTTTACCGTTGCTC

-520 CAAACCTCCCCTCCATTTTTTAATGGCGGTATTAGAGTCCGATGGGAGGG

-470 CCACACTCAAAGGAAATGATGTACACAGTTTATCCTTATGCACGTAGTAA

-420 TGACTGCTTTCACGGCTCAAACTCATGACCTATAAGTCACACATAGACAA

-370 TTTTACCGGTGCTCAAGCTCCCTCATTCTTTTAATTACAGAGCACATTAA

SRE box

-320 CCATAATTGAGCATAGATGAGGCACTACTGATTGTCTTCATCACGGTATT

-270 TTATATCAGTACGACATGATGAAAGCACCAAGACTAAATACAGGGGTAAA

-220 ATTTAGACTAGGCCTTTACTTTTTAGAAGCCTAGGAAGCTCACACTTGCG

-170 AGTGGAAAAGAACACAAGGCCAAACCCAATAGGGAAAATGCCAAATTTGA

TATA box

-120 TGGTGGCAATTAGCAATTATCTGTATTTGATTATAACACCAAAACAAGAA

RB

-70 GAAAGAGTAATAAACAAATTATTCTAATTCGTTTACCCTCACACTGATAG

BamHI SalI +1

-20 TTTCGGATCCGGTCGACTCTATG(The start codon of the *lux A* gene)

Fig. V-6. Comparison of the T-DNA right border sequence before and after integration. The T-DNA right border (25 bp) sequence is bold and underlined. (**a**) Before T-DNA was integrated into the plant genome. (**b**) After T-DNA was integrated into plant genome.

(a) 5'-**GTTTACCCGCCAATATATCCTGTCA**AACACTGATAGTTTC-3'

(b) 5'-**GTTTACCC**.....**TCA**.....CACTGATAGTTTC-3'

Fig. V-7. Construction of a plant promoter test vector, pTV-1p, containing the FEP-*lux F* fusion. (**a**) Donor plasmid: A recombinant clone, pBS-1, containing a full length insert of the putative plant promoter, FEP. pBS-1 is derived from pBluescript KS after insertion of the FEP DNA fragment. MCS, multiple cloning sites; f1(+/-)ori, f1 origin; LacZ and Lac1, *Lac Z* gene of *E. coli*; Co1E1ori, Co1E1 origin. The FEP DNA fragment in pBS-1 was released by Xba I \ Sal I double digestions and was cloned into a modified binary plant transformation vector, pPCV 701 *luxFM2* , to replace the mas promoter (P1 and P2). This formed the plant promoter test vector, pTV-1p. The T-DNA between the right and left borders (RB and LB) indicates the transferred DNA which is stably integrated into the plant genome. The arrows indicate the direction of transcription; P_{NOS} is the nopaline synthase promoter, Pg5 is the promoter of the T_L-DNA gene 5; g4pA, OcspA, and g7pA are polyadenylation sequences derived from T_L-DNA gene 4, the octopine synthase gene, and gene 7; NPTII is the neomycin phosphotransferase gene; ori_T and ori_V are replication and conjugational-transfer origin sequences derived from plasmid pRK2, oripBR is the replication of origin of pBR322 and Ap^r is the β-lactamase gene from *E. coli*. (**b**) Junction region of the FEP-*lux F* fusion. The ATG is the start codon of the *lux F* gene. Restriction enzyme sites are marked and underlined.

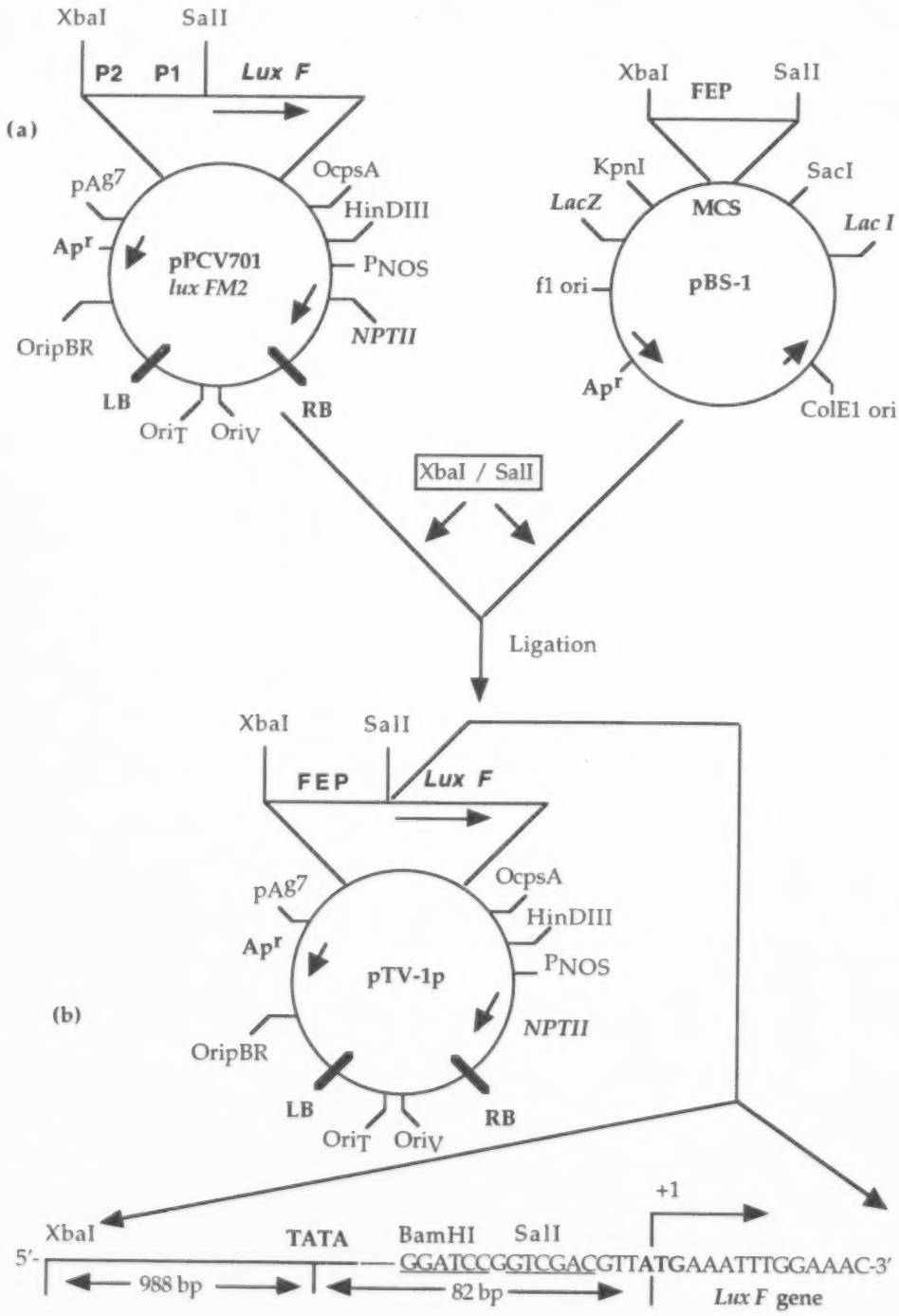
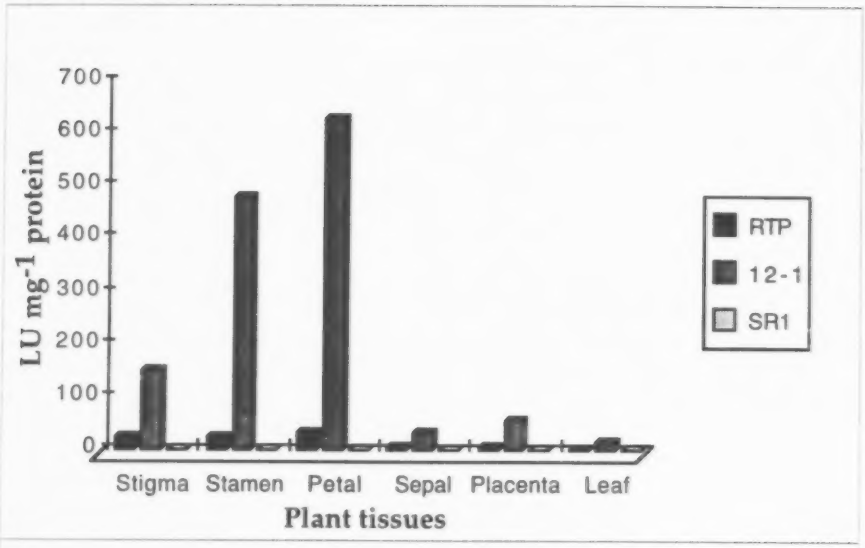
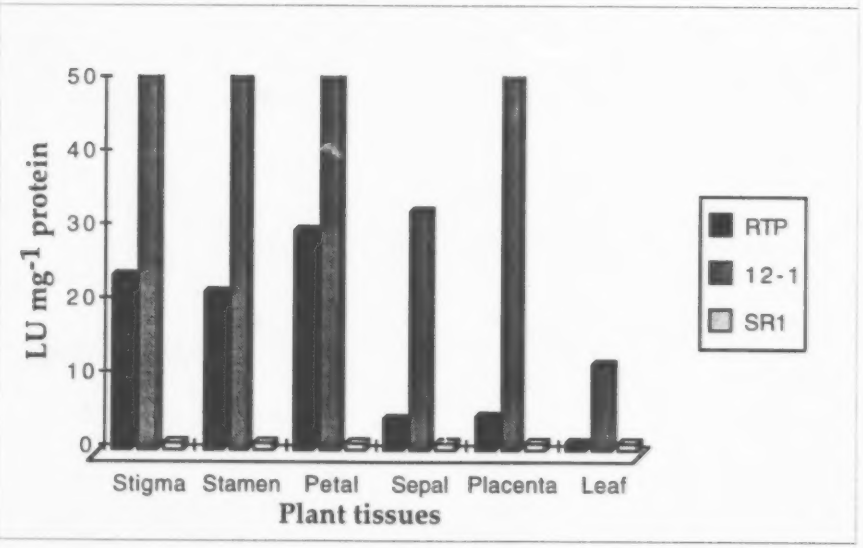

(a)
XbaI
SalI
P2
P1
Lux F
pAg7
OcpsA
HinDIII
Ap^r
pPCV701
lux FM2
PNOS
NPTII
OripBR
LB
RB
OriT
OriV
XbaI
SalI
FEP
KpnI
SacI
LacZ
MCS
Lac I
f1 ori
pBS-1
Ap^r
ColE1 ori
XbaI / SalI
Ligation
(b)
XbaI
SalI
FEP
Lux F
pAg7
OcpsA
HinDIII
Ap^r
pTV-1p
PNOS
NPTII
OripBR
LB
RB
OriT
OriV
+1
XbaI
TATA
BamHI
SalI
5'-
GGATCCGGTCGACGTTATGAAATTTGGAAAC-3'
988 bp
82 bp
Lux F gene

Chart V-1. FEP-directed *lux* gene expression in floral organs of transgenic tobacco plants. Different floral tissues (stigma, stamen, petal, sepal and placenta) and leaf tissues from retransformed tobacco plants (RTP) containing the promoter test vector, pTV-1p, along with plant 12-1 and SR1 (nontransgenic plant as a control, CK) were screened for luciferase activity by a luminometer. (**a**) Luciferase activity detected in different floral organs. (**b**) The same chart as (**a**), but with smaller scale. The value in RTP represents the average of three assays for each tissue with about 50 mg fresh weight each from nine independent transgenic plants. The value in plant 12-1 and SR1 represents the average of three assays for each tissue with about 50 mg fresh weight from three individual plants. Luciferase activities are expressed as LU mg $^{-1}$ protein where 1 LU is equal to 1.2 x 10^6 protons sec^{-1}.

a.


700
600
500
400
300
200
100
0
LU mg-1 protein
RTP
12-1
SR1
Stigma
Stamen
Petal
Sepal
Placenta
Leaf
Plant tissues

b.


50
40
30
20
10
0
LU mg-1 protein
RTP
12-1
SR1
Stigma
Stamen
Petal
Sepal
Placenta
Leaf
Plant tissues

Chapter VI

General Discussion and Future Perspectives

GENERAL DISCUSSION

In the present study, the bacterial luciferase genes from a marine bacterium *Vibrio harveyi* (Belas *et al.*, 1982; Baldwin *et al.*, 1984; Olsson *et al.*, 1989) were used as marker genes to modified a binary plant transformation vector and to develop a T-DNA tagging system in higher plants. As discussed in Chapter III, two new vectors, pPCV701 *luxFM1* and pPCV701 *luxFM2* , were constructed from the vector, pPCV701 *lux F*, keeping the original features (Koncz *et al.*, 1987) while improving the cloning capabilities and expression of desired genes in higher plants. Since *Agrobacterium-mediated* transformation is the most frequently used method to introduce foreign genes into plant cells and to create transgenic plants (Schell, 1987; Gasser and Fraley, 1989), a convenient binary plant transformation vector provides a powerful tool for study of gene structure, function and expression. Removal of internal identical restriction enzyme sites in the vector allows us to clone a gene of choice in a single step. Insertion of multiple cloning sites in the vector offers more choices for cloning strategies (Jiang *et al.*, 1993a).

As discussed in Chapter IV, direct introduction of a promoterless *lux* marker gene into a plant genome to monitor the activation of a plant gene *in vivo* offers a novel method to study gene expression in the native environment. Compared with the promoterless GUS gene fusion system previously reported (Jefferson *et al.*, 1987; Kertbundit *et al.*, 1991) or the promoterless *aph(3')-II* gene fusion system (Koncz *et al.*, 1989), the promoterless *lux* gene fusion system provides a convenient tool to monitor temporal and spatial patterns of gene expression during plant development. The promoterless bacterial luciferase marker gene is part of a functional insertion element integrated into the plant genome, which brings about the *in vivo* expression of the luciferase marker gene in transgenic tobacco plants.

This can be analyzed by low-light video imaging (Jiang *et al.*, 1992; 1993b). Therefore, identification of insertional mutants can be based on marker gene expression rather than on phenotype, a major criticism of other T-DNA tagging methods (Koncz *et al.*, 1992). This feature also permits T-DNA tagging of polyploid plant species.

The major concern in all T-DNA tagging methods is their low mutation frequency (Walden *et al.*, 1991). According to a previous report, (Feldmann, 1991), if mutant selection is based on phenotype, generation of a completely saturated population of transformants i.e. a 95% probability of an insert in any 2 kb of DNA, requires 75,000 individual transformants in *Arabidopsis*. The results from this experiment indicate that selection based on the luciferase gene expression can increase the frequency. Combining the iPCR method (Ochman *et al.*, 1988) to facilitate amplification of flanking plant DNA fragments further increases the efficiency of this approach (Jiang *et al.*, 1992).

T-DNA tagging experiments with the promoterless luciferase gene yielded a large number of insertional mutant plants (45) that showed various expression patterns of the luciferase gene, such as root-, stem- and flower-specific gene expression. In Chapter V, one insertional mutant (plant 12-1) which showed the luciferase gene expression predominantly in the flower was studied in detail. DNA sequence analysis of a putative promoter element that controlled the tissue specific gene expression revealed TATA-like box and CAAT boxes located at proximal region adjacent to the promoterless *lux A* gene. In addition, a consensus sequence similar to the control region of other flower specific genes was identified (Schwarz-Sommer *et al.*, 1992). When this putative promoter region was fused with the *lux F* reporter gene and reintroduced into tobacco, luciferase gene expression was detected in

flowers of the transgenic tobacco. These results confirm that the putative plant promoter is responsible for mediating tissue-specific expression of the reporter gene in transgenic tobacco plants. These results also suggests that the use of the promoterless bacterial luciferase gene to monitor the activation of a plant gene *in vivo* before its isolation is a realistic approach to isolate promoters from higher plants. Therefore, it should be worthwhile to continue to study these mutant plants individually at the molecular level.

In conclusion, the present study is the first time the use of the bacterial luciferase gene in T-DNA tagging experiments has been documented. Combined with iPCR, this approach has proven to be a simple, sensitive and reliable method for studying regulation of gene expression, insertional activation and isolation of promoter elements from higher plants. Furthermore, the construction of luciferase gene-based binary plant transformation vectors provides a convenient tool for cloning and expression of a foreign gene in higher plants.

FUTURE PERSPECTIVES

T-DNA tagging has proven to be a very powerful method for identifying a wide range of mutants (Feldmann, 1991; Walden *et al.*, 1991). Once insertional mutants are created, T-DNA becomes a permanent molecular tag in the genome (Walbot, 1992). The putative promoter element identified in the present study can be used directly as a hybridization probe to isolate the corresponding gene from a tobacco genomic library or in RFLP experiments to identify a nearby gene by chromosome walking (Chang *et al.*, 1988).

It is also worthwhile to isolate more promoter elements from those insertional mutants obtained from the present study in order to further

understand gene regulation with regard plant development. Isolation and characterization of native promoters from higher plants will, over the long term, assist us to use genetic engineering as a tool for crop improvement.

So far, the most frequently used promoters in crop improvement are bacterial or viral promoters such as bidirectional mas promoters from mannopine synthase (Velten *et al.*, 1984; Koncz *et al.*, 1987; Jiang *et al.*, 1993a), or the nos promoter from nopaline synthase of T-DNA genes (Depicker *et al.*, 1982), or the CaMV 35S RNA promoter from the cauliflower mosaic virus (Odell *et al.*, 1985; Banfey and Chua, 1989). The applications of these bacterial or viral promoters are limited by their regulatory patterns. Thus, isolation of plant promoters with various regulatory roles will meet the increasing demands of biotechnology.

T-DNA tagging associated with other methods will generate many novel applications. For example, T-DNA tagging vectors carrying a gene with a weak (or a truncated) promoter in the wrong orientation, capable of creating a small number of antisence transcripts (not enough to shut down the gene), can be used to study gene function. Insertion of this construct into a region associated with a temporal enhancer may retard gene expression at a stage or in an organ. This type of approach will provide more information on the gene's involvement in plant development. Many types of genes can be used in this approach, such as a ribonuclease gene (Mariani *et al.*, 1990), a cell cycle gene (*cdc2*, cyclin A and B genes) (Doonan, 1991), or a gene involved in development during cytodifferentiation and organogensis.

The principal of the bacterial luciferase based T-DNA tagging system in higher plants can be applied to other systems and other organisms. The promoterless bacterial luciferase gene can be linked to a transposon and used directly in transposon tagging in maize, snapdragon and *Arabidopsis;*

(Walbot, 1992), or to a P element in yeast and *Drosophila* (Roeder *et al.*, 1980; Bingham *et al.*, 1981), or to a retrovirus for transfection of mammalian cells (Gridley *et al.*, 1987). So far these types of applications have not been reported. There is no doubt that research in this direction will significantly increase our knowledge of gene structure, function and regulation in higher plants and animals.

BIBLIOGRAPHY

Baldwin, T.O., Berends, T., Bunch, T.A., Holzman, T.F., Rausch, S.K., Shamansky, L., Treat, M. and Ziegler, M.L. (1984) Cloning of the luciferase structural genes from *Vibrio harveyi* and expression of bioluminescence in *Escherichia coli*. *Biochemistry* 23, 3663-67.

Belas, R., Mileham, A., Cohn, D., Hilmen, M., Simon, M. and Silverman, M. (1982) Bacterial bioluminescence: Isolation and expression of the luciferase genes from *Vibrio harveyi*. *Science* 218, 791-92.

Benfey, P.N. and Chua, N.H. (1989) Regulated genes in transgenic plants. *Science* 242, 174-81.

Bingham, P., Lewis, R. and Rubin, G.M. (1981) Cloning of DNA sequences from the white locus of *D. melanogaster* by a novel and general method. *Cell* 20, 579-95.

Chang, C., Bowman, J.L., Dejohn, A.W., Lander, E.S. and Meyerowitz, E.M. (1988) Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 85, 68-6860.

Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. and Goodman, H.M. (1982) Nopaline synthase: transcript mapping and DNA sequence. *J. Mol. Appl. Gen.* 1, 561-73.

Doonan, J.H. (1991) Cycling in plant cells. *Plant J.* 1, 129-32.

Feldmann, K.A. (1991) T-DNA insertion in *Arabidopsis*: mutational spectrum. *Plant J.* 1, 71-82.

Gasser, C.S. and Fraley, R.T. (1989) Genetically engineering plants for crop improvement. *Science* 244, 1293-99.

Gridley, T., Soriano, P. and Jaenisch, R. (1987) Insertional mutagenesis in mice. *Trends Genet.* 3, 162-66.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M. (1987) GUS fusion: β–glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6, 3901-7.

Jiang, C., Escher, A. and Szalay, A.A. (1993a). Alteration of a luciferase based binary plant transformation vector for simultaneous expression of desired gene products and bacterial luciferase in transgenic plants. Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds.) John Wiley & Sons. 127-31.

Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1992) Identification of plant genes *in vivo* by tagging with T-DNA border-linked luciferase genes followed by inverse polymerase chain reaction amplification. *Plant Mol. Biol. Rep.* 10, 345-61.

Jiang, C., Langridge, W.H.R. and Szalay, A.A. (1993b) Isolation of developmentally regulated promoter elements from tobacco by T-DNA insertional activation with promoterless luciferase marker gene. Bioluminescence and Chemiluminescence, Status Report. Szalay, A.A., Kricka, L. and Stanley, P. (Eds.) John Wiley & Sons. 217-21

Kertbundit, S., De Greve, H., DeBoeck, F., Van Montagu, M. and Hernalsteens, J.P. (1991) *In vivo* random β-glucuronidase gene fusion in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 88, 5212-16.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Korber, H., Redei, G.P. and Schell, J. (1989) High-frequency T-DNA-mediated gene tagging in plants. *Proc. Natl. Acad. Sci. USA* 86, 8467-71.

Koncz, C., Olsson, O., Langridge, W.H.R., Schell, J. and Szalay, A.A. (1987) Expression and assembly of functional bacterial luciferase in plants. *Proc. Natl. Acad. Sci. USA* 84, 131-35.

Koncz, C., Schell, J. and Redei, G.P. (1992) T-DNA transformation and insertion mutagenesis. Methods in *Arabidopsis* Research. Koncz, C., Chua, N.H. and Schell, J. (Eds.) World Scientific. 224-73.

Mariani, C., De Beuckeleer, M., Truettner, J., Leemans, J. and Goldberg, R.B. (1990) Induction of male sterility in plants by a chimeric ribonuclease gene. *Nature* 347, 737-41.

Ochman, H., Gerber, A.S. and Hartl, L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* 120, 621-23.

Odell, J.T., Nagy, F. and Chua, N.H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic 35S promoter. *Nature* 313, 810-12.

Olsson, O., Escher, A.P., Sandberg, G., Schell, J., Koncz, C. and Szalay, A.A. (1989) Engineering of monomeric bacterial luciferases by fusion of *luxA* and *luxB* genes in *Vibrio harveyi*. *Gene* 81, 335-47.

Roeder, G.S., Farabaugh, P.J., Chaleff, D.F. and Fink, G.R. (1980) The origins of gene instability in yeast. *Science* 209, 1375-80.

Schell, J. (1987) Transgenic plants as tools to study the molecular organization of plant genes. *Science* 237, 1176-82.

Schwarz-Sommer, Z., Hue, I., Huijser, P., Flor, P.J., Hansen, R., Tetens, F., Lonnig, W.E., Saedler, H. and Sommer, H. (1992) Characterization of the *Antirrhinum* floral homeotic MADS-box gene *DEFICIENS*: Evidence for DNA binding and autoregulation of its persistent expression throughout flower development. *EMBO J.* 11, 251-63.

Velten, J., Velten, L., Hain, R. and Schell, J. (1984) Isolation of a dual plant promoter fragment from the Ti plasmid of *Agrobacterium tumefaciens*. *EMBO J.* 3, 2723-30.

Walbot, V. (1992) Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis. *Annu. Rev. Plant Phhysiol. Plant Mol. Biol.* 43, 49-82.

Walden, R., Hayashi, H. and Schell, J. (1991) T-DNA as a gene tag. *Plant J.* 1, 281-88.

Walden, R., Koncz, C. and Schell, J. (1990) The use of gene vectors in plant molecular biology. *Methods Mol. Cell Biol.* 1, 175-94.